I segreti dei Supervulcani

Microscopici cristalli di cenere vulcanica stanno fornendo informazioni sorprendenti sulle eruzioni più devastanti del mondo

di Ilya N. Bindeman

ANELLO DI FUOCO. Crateri grandi come una montagna esplodono attorno al bordo esterno di un supervulcano attivo e ricoprono il paesaggio di nubi di gas e di ceneri roventi.

Julia Green

Nascosti in profondità sotto la superficie della California e del Wyoming ci sono due vulcani in letargo, potenzialmente capaci di una violenza inimmaginabile. Se mai dovessero entrare in fase critica, potrebbero seppellire in poche ore tutta la parte occidentale degli Stati Uniti sotto due metri di ceneri; lo hanno già fatto per almeno quattro volte negli ultimi due milioni di anni. Supervulcani simili si trovano anche al di sotto dell'Indonesia e della Nuova Zelanda. L'eruzione di un supervulcano ha la forza devastante di un piccolo asteroide in collisione con la Terra e si verifica statisticamente dieci volte più spesso: queste esplosioni sono una delle catastrofi naturali più tremende che l'umanità possa subire. Oltre alla distruzione immediata provocata dalla ricaduta di ceneri roventi, i supervulcani attivi sprigionano gas che alterano radicalmente il clima globale per anni e anni.

Ovviamente i ricercatori non vedono l'ora di capire che cosa può scatenare la fase eruttiva di questi giganti in sonno, come si può prevedere la ripresa della loro furia distruttiva e quali possono essere le conseguenze di uno di questi eventi. Qualche risposta è arrivata dalle recenti analisi di microscopici cristalli contenuti nei depositi di ceneri delle loro antiche eruzioni. Queste osservazioni, oltre alla disponibilità di tecnologie più potenti per il monitoraggio dei potenziali siti di disastro, fanno sperare nella possibilità, per il futuro, di individuare con largo anticipo i segni premonitori dell'entrata in attività di questi vulcani. Altri studi suggeriscono però che le emissioni di un supervulcano sono in grado innescare pericolose reazioni chimiche nell'atmosfera, in grado di rendere i mesi successivi all'eruzione ancora più micidiali di quanto si pensasse fino a oggi.

Quasi tutti i vulcanologi ritengono che gli attuali abitanti della Terra abbiano una probabilità estremamente remota di vivere l'eruzione di un supervulcano. Le eruzioni catastrofiche tendono a verificarsi soltanto una volta ogni centinaia di migliaia di anni. Tuttavia, data l'enormità delle dimensioni e degli effetti globali di questi episodi, l'attività dei supervulcani è oggetto di studio fin dagli anni cinquanta.

Terrore primordiale

Una delle prime scoperte dei geologi è stata l'esistenza di enormi valli circolari – alcune di diametro compreso fra 30 e 60 chilometri e della profondità di diversi chilometri – che apparivano molto simili alle caldere a forma di ciotola situate alla sommità di molti fra i più noti vulcani del pianeta. Normalmente le caldere si formano quando la camera magmatica contenente la roccia fusa che si trova al di sotto di un camino vulcanico si svuota, provocando il collasso del terreno sovrastante. Notando che le valli a forma di caldera si trovavano vicino ad alcuni dei più imponenti depositi di rocce vulcaniche generati nel corso di un solo evento, i ricercatori si resero conto di trovarsi di fronte ai resti di vulcani centinaia o anche migliaia di volte più imponenti del famigerato Mount St. Helens, nello Stato di Washington. Dalle immani dimensioni delle caldere e dal volume di materiale eruttato, si poteva desumere che le camere magmatiche dovevano essere altrettanto mostruose.

Visto che i fenomeni di generazione di calore in grado di formare camere magmatiche così vaste sono rari, i supervulcani sono altrettanto infrequenti. Negli ultimi due milioni di anni vi sono state solo quattro esplosioni con un'emissione minima di 750 chilometri cubi di materiale tutto in una volta, in quattro regioni: lo Yellowstone National Park nel Wyoming, la Long Valley in California, Toba a Sumatra e Taupo in Nuova Zelanda. La ricerca dei resti di eruzioni altrettanto vaste prosegue in altre aree di crosta continentale spessa, tra cui il Sud America occidentale e l'estremo lembo orientale della Russia.

A metà degli anni settanta gli studi hanno rivelato alcuni modi in cui la camera magmatica si forma e diventa pericolosa. Sotto la superficie di Yellowstone, la placca tettonica nordamericana sta spostandosi al di sopra di un pennacchio di roccia fusa viscosa che va risalendo attraverso il mantello, lo strato interno della Terra spesso 2900 chilometri che si trova tra il nucleo fuso e quell'involucro relativamente sottile che è la crosta. Funzionando come un colossale becco di Bunsen, questo cosiddetto *hot spot* ha provocato la fusione della crosta sovrastante in quantità tali da alimentare eruzioni catastrofiche per gli ultimi 16 milioni di anni. A Toba l'origine della camera è differente. Quella regione si trova infatti al di sopra di una zona di subduzione, un'area in cui una zolla tettonica sta scorrendo al di sotto di un'altra; la convergenza produce un'esteso riscaldamento, principalmente attraverso la fusione parziale del mantello al di sopra della zolla che sta affondando.

Indipendentemente dall'origine del calore, la pressione all'interno della camera magmatica aumenta nel tempo, via via che una maggiore quantità di magma si accumula sotto il peso enorme delle rocce sovrastanti. Una supereruzione si verifica quando il magma sotto pressione crea un rigonfiamento nella crosta che lo ricopre, sufficiente a produrre fratture con andamento verticale che giungono fino alla superficie. Il magma risale poi lungo queste nuove spaccature, e finisce con il formare un anello di bocche vulcaniche di eruzione. Quando le bocche si saldano le une con le altre, l'enorme cilindro di crosta all'interno dell'anello non ha più nulla che lo possa sorreggere.

Questa «volta» di solida roccia precipita allora verso il basso su ciò che è rimasto del magma sottostante. Il collasso espelle violentemente la quantità restante di magma e di gas attorno ai bordi dell'anello (*si veda il box a pp. 46-47*).

In sintesi/*Immani eruzioni*

- Le recenti analisi sulla composizione di microscopici cristalli all'interno di depositi di ceneri vulcaniche di eruzioni preistoriche stanno ribaltando le convinzioni sul comportamento dei supervulcani e svelano alcune delle sorprendenti conseguenze immediate della loro attività eruttiva.
- Il funzionamento interno delle camere magmatiche che alimentano i supervulcani potrebbe evolvere in modo da influenzare fortemente le eruzioni future.
- L'inverno vulcanico che investirebbe l'intero pianeta dopo una supereruzione sarebbe probabilmente più breve del previsto, ma alcune reazioni chimiche nell'atmosfera potrebbero rivelarsi molto più pericolose.

GRANDE, PIÙ GRANDE, GRANDISSIMO

Poiché espellono una quantità di materiale assai maggiore e con una violenza molto più accentuata, i supervulcani (*arancione e blu*) spargono ceneri a distanze ben più grandi rispetto ai vulcani «normali» (*giallo e viola*), anche durante le eruzioni di maggiore intensità.

Jen Christiansen; fonte U.S. Geological Survey

Le «impronte digitali»

Restavano però alcuni misteri. In particolare, come compresero presto i ricercatori, non tutte le grandi camere magmatiche finiscono con il generare un'eruzione catastrofica. A Yellowstone, per esempio, ci sono tre delle più giovani caldere di supervulcano al mondo (formatesi 2,1 milioni, 1,3 milioni e 640.000 anni fa, quasi una sopra l'altra), ma negli intervalli tra questi eventi esplosivi la camera sottostante ha rilasciato volumi di magma simili in modo lento e tranquillo. La ragione per cui il magma a volte riesce a diffondere lentamente fino alla superficie è ancora poco chiara.

Uno sguardo alla composizione dei minuscoli cristalli intrappolati all'interno della lava e della cenere eruttate a Yellowstone ha suggerito una parte della risposta, offrendo nuove informazioni sul modo in cui si forma il magma. Per decenni i geologi hanno supposto che il magma stazionasse per milioni di anni e che in occasione di ogni eruzione nuova roccia allo stato liquido risalisse per ristabilire immediatamente il contenuto della camera magmatica. Se l'ipotesi fosse corretta, le grandi eruzioni catastrofiche dovrebbero essere molto più numerose, perché è impossibile da un punto di vista termico e meccanico mantenere ammassi di magma di enormi dimensioni all'interno della crosta senza svuotarli frequentemente.

La vecchia idea si basava in gran parte sulle analisi di roccia intera, in cui i ricercatori ottenevano un unico insieme di parametri chimici per ciascun campione – delle dimensioni di un pugno – di roccia formata da ceneri vulcaniche. Quei dati fornivano importanti modelli generali dell'evoluzione del magma, ma erano insufficienti per determinare l'età del magma eruttato e la profondità della sua formazione.

Ogni frammento di roccia è costituito da migliaia di minuscoli cristalli, ognuno con composizione e storia peculiari. Quando alla fine degli anni ottanta è stata possibile un'analisi più precisa dei singoli cristalli, è stato come riuscire leggere tutti i capitoli di un libro anziché basarsi sul risvolto di copertina. Per esempio si è scoperto che alcuni cristalli, e quindi i magmi in cui si sono formati, si sono originati molto prima di altri, e che alcuni si sono formati a grande profondità, mentre altri in prossimità della superficie.

Negli ultimi dieci anni i geochimici hanno rivolto l'attenzione a un cristallo vulcanico particolarmente durevole: lo zircone. Sapendo che i cristalli di zircone possono sopportare variazioni di temperatura e pressione estreme senza alterare la propria composizione originale, alcuni ricercatori li hanno usati per studiare le prime fasi di evoluzione della crosta terrestre (si veda *Il grande freddo della giovane Terra*, di John W. Valley, in «Le Scienze» n. 448, dicembre 2005). Quando mi aggregai al gruppo di Valley, nel 1998, usammo gli zirconi di Yellowstone per ricostruire la storia del magma che li aveva generati, scoprendo importanti indizi su come il vulcano potrebbe comportarsi in futuro.

Il primo passo fu misurare i rapporti di diverse forme isotopiche dell'ossigeno in zirconi derivanti dalla più recente supereruzione di Yellowstone, che 640.000 anni fa diede luogo alla formazione dei tufi di Lava Creek, un deposito di ceneri vulcaniche che in alcune località raggiunge

Cortesia J.S. Lackey/College of Wooster; dati U.S. Geological Survey (*in alto*); Todd Cullings/National Park Service (*in basso*)

I SUPERVULCANI QUIESCENTI, come quello della Long Valley in California (*in alto*), non hanno la consueta forma a cono del Mount St. Helens (*in basso*), nello Stato di Washington. Sono invece caratterizzati da enormi caldere, ovvero depressioni della crosta terrestre formatesi con il collasso del terreno all'interno delle camere magmatiche che si erano svuotate alimentando le supereruzioni più recenti.

no formati da rocce che hanno interagito con pioggia o acqua di mare. Sospettavamo quindi che la roccia del tetto di camera magmatica collassato in una delle due più antiche supereruzioni di Yellowstone dovesse essersi fusa per formare il grosso del magma espulso durante la più recente catastrofe di Lava Creek e le piccole eruzioni avvenute in seguito. Questa ipotesi ha acquistato valore nel momento in cui abbiamo appurato che le età degli zirconi derivanti dalle eruzioni successive all'episodio di Lava Creek coprono l'intera durata di due milioni di anni del vulcanismo di Yellowstone. Zirconi così antichi potrebbero esistere nella cenere più giovane solo se fosse stato il magma originale ad alimentare l'intera serie delle eruzioni.

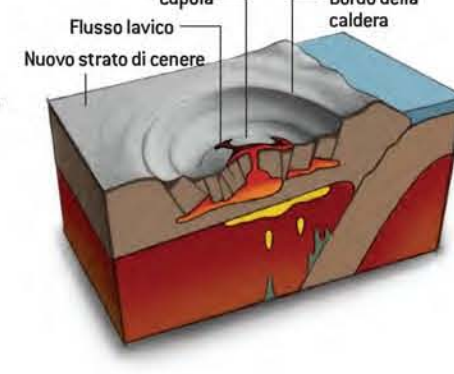
La nostra scoperta implica che oggi si può pensare di fare alcune previsioni su come il supervulcano di Yellowstone, e magari anche quelli individuati altrove, potranno comportarsi in futuro. Se una nuova scarica di piccole eruzioni dovesse avere inizio a Yellowstone, come accade un po' prima di un'eruzione catastrofica (un «prima» che può variare da alcune settimane a qualche secolo), la misura dei rapporti isotopici dell'ossigeno di quelle lave e la determinazione dell'età degli zirconi in esse contenuti dovrebbero rivelare quale tipo di magma abbonda nella camera sottostante. Se la prossima eruzione dovesse rivelarsi povera di ossigeno-18, ciò segnalerebbe il perdurare della sua alimentazione con residui stagnanti del magma originale, ormai più simile a una spessa poltiglia cristallina che non a un liquido esplosivo. Invece se la nuova lava presentasse i rapporti isotopici caratteristici di nuovo magma proveniente dal mantello, e non contenesse vecchi zirconi, molto probabilmente deriverebbe da un grande volume di nuovo magma arrivato dal basso per riempire nuovamente la camera magmatica. Questi riscontri implicherebbero che ha avuto inizio un nuovo ciclo di vulcanismo, e indicherebbero un maggiore potenziale della camera di esplodere in modo catastrofico.

SUPERCICLI PER SUPERVULCANI

Le grandi camere di magma fuso che alimentano i supervulcani si formano al di sopra di punti caldi (gli *hot spot*, pennacchi di magma che risalgono dalle profondità del mantello) e di zone di subduzione (regioni in cui una placca tettonica sta scorrendo sotto un'altra). In entrambi i casi, i vulcani giganti tendono a seguire un ciclo eruttivo che può essere più prevedibile di quanto si pensasse in passato. Ecco i quattro passaggi di base, a partire dalla formazione iniziale della camera magmatica, rappresentati per una zona di subduzione.

1 La fusione parziale di roccia del mantello sopra la placca di crosta oceanica in affondamento produce magma che si fa strada verso la base della crosta continentale, dove si raccoglie. Questa camera magmatica inferiore funziona come un enorme becco Bunsen che finisce per fondere parti della crosta continentale, che ha un punto di fusione inferiore a quello della roccia sottostante. Parte del magma risale anche lungo piccoli condotti verticali tra le due camere.

2 Al riempirsi della camera magmatica, il terreno sovrastante si rigonfia e si spacca. Questo magma è ricco di silice, e ha una temperatura più bassa rispetto al magma che si forma nel mantello, è resistente allo scorrimento e sia l'acqua che il gas hanno difficoltà a risalirlo. Quindi, quando una protrusione di questo magma viscoso si fa strada fino alla superficie il materiale sottostante ad alta pressione tende a esplodere con violenza anziché scorrere lentamente.

3 La superficie terrestre sottoposta a sollecitazione collassa, mentre nuove bocche esplosive formano un anello ampio come la camera magmatica. I pezzi di roccia fratturata cadono nella camera, forzando altro magma a risalire lungo i bordi dell'anello. Rilasciato all'improvviso, il magma si trasforma in enormi e incandescenti nubi di ceneri, gas e roccia: i cosiddetti flussi piroclastici, o nubi ardenti, che distruggono il paesaggio nel raggio di decine di chilometri.

4 In seguito all'eruzione, una depressione a forma di cratere, la caldera, insiste su una camera magmatica parzialmente vuota. Con il tempo, il terreno collassato all'interno della camera inizia a fondere dall'interno, generando nuovo magma che insieme ad altre forze provoca la formazione di un rigonfiamento a cupola nel centro della caldera. Da questa zona può disperdersi a più riprese lava in lento movimento, prima che si accumuli magma a sufficienza per alimentare una nuova supereruzione.

Jen Christiansen

uno spessore di 400 metri, nonché in zirconi contenuti in depositi ancora più recenti, espulsi da eruzioni di minore entità. Terminate le analisi iniziali constatammo con sorpresa che la composizione dell'ossigeno degli zirconi non corrispondeva a quella del mantello caldo e profondo, come dovrebbe essere se le camere magmatiche, una volta svuotate, venissero costantemente rifornite dagli strati sottostanti. Gli zirconi provenienti da magmi derivanti dal mantello avrebbero dovuto mostrare una «firma» distintiva: quando gli elementi chimici che sono disciolti nei magmi si combinano a formare uno zircone, quel cristallo include una frazione elevata di ossigeno-18, un isotopo dell'ossigeno che contiene dieci neutroni.

Capimmo subito che il magma doveva essersi originato in rocce un tempo vicine alla superficie terrestre. Gli zirconi da noi studiati, infatti, erano poveri di ossigeno-18 rispetto al mantello, e questo impoverimento si verifica solo se i cristalli si so-

Dopo l'eruzione

I minuscoli cristalli e i loro rapporti isotopici hanno anche rivelato sorprendenti dettagli sui periodi immediatamente successivi alle supereruzioni. Uno degli esempi meglio studiati delle conseguenze di attività supervulcanica è la formazione dei tufi di Bishop, uno strato vulcanico spesso anche qualche centinaio di metri che affiora nelle Volcanic Tablelands della California orientale. Questo imponente deposito è ciò che rimane dei 750 chilometri cubi di magma che secondo le stime furono espulsi durante la formazione della caldera del supervulcano di Long Valley, circa 760.000 anni fa.

Prima degli anni settanta, molti geologi supponevano che per produrre i tufi di Bishop dovesse essersi verificata una serie di eruzioni distinte nell'arco di diversi milioni di anni. Ma attenti studi condotti su microscopiche bolle piene di magma intrappolate all'interno di minuscoli cristalli di quarzo hanno indotto a rivedere la storia. La velocità con cui il magma lascia una camera magmatica dipende in primo luogo da due fattori: la viscosità del magma e la differenza di pressione tra la camera e la superficie terrestre. Poiché la pressione all'interno di una bolla corrisponde a quella della camera in cui quel magma si è formato, la bolla rappresenta una versione miniaturizzata della camera stessa. Coscienti di questa corrispondenza, Alfred Anderson dell'Università di Chicago e i suoi colleghi hanno punto le bolle sotto un microscopio, osservando quanto tempo impiegava il magma a fuoriuscire.

Basandosi su questi e altri esperimenti effettuati negli anni novanta, oggi i geologi ritengono che i tufi di Bishop, e probabilmente la maggior parte degli altri detriti di supereruzione, siano stati espulsi in un unico evento durato tra 10 e 100 ore.

A partire da quella scoperta, i ricercatori hanno dovuto modificare le loro ricostruzioni delle eruzioni supervulcaniche. Ecco che cosa si aspettano ora da un evento della scala di quelli che hanno colpito Long Valley e Yellowstone: anziché un lento flusso di lava incandescente come quella che si vede scorrere lungo i versanti del vulcano Kilauea, nell'isola di Hawaii, queste eruzioni darebbero luogo a esplosioni supersoniche di una miscela surriscaldata dall'aspetto schiumoso di gas e ceneri che verrebbe proiettata nella stratosfera fino a 50 chilometri di altezza. Contemporaneamente al collasso del terreno al di sopra della camera magmatica, immensi flussi piroclastici dall'aspetto di dense nubi grigie si sprigionerebbero in modo esplosivo tutto attorno alla caldera. Questi flussi sono uno stadio intermedio tra la lava e la cenere, e si spostano in modo rapidissimo: secondo alcune fonti fino a 400 chilometri all'ora; nessuna automobile avrebbe alcuna possibilità di batterli in velocità. I flussi piroclastici, o nubi ardenti, sono estremamente caldi, da 600 a 700 gradi, e quindi bruciano tutto ciò che incontrano sul loro tragitto, oltre a seppellirlo.

Ma se i flussi piroclastici sono micidiali, le ceneri immesse nell'atmosfera possono avere conseguenze ancora più catastrofiche. Nel raggio di centinaia di chilometri attorno all'eruzione, e per giorni o forse settimane, una cenere di colore grigio chiaro cadrebbe quasi fosse neve. Entro 200 chilometri di distanza dalla caldera, la maggior parte della luce solare verrebbe schermata, e il cielo di mezzogiorno sembrerebbe come quello del crepuscolo. Case, persone e animali sarebbero sepolti, e in qualche caso schiacciati. Ancora a 300 chilometri di distanza la cenere potrebbe raggiungere lo spessore di un metro: mescolata alla pioggia acquisirebbe un peso sufficiente a schiantare i tetti delle case. Una quantità di cenere anche minore basterebbe a mettere fuori uso le centrali elettriche e i ripetitori. E uno strato di copertura di un solo millimetro, quale si potrebbe depositare sull'intero emisfero avente al proprio centro il supervulcano, basterebbe a provocare la chiusura degli aeroporti e a ridurre drammaticamente la produzione agricola.

Solo gradualmente la pioggia potrebbe lavar via la spessa coltre di ceneri, e da-

U.S. Department of Energy

I DUE DEPOSITI VULCANICI visibili lungo un ripido pendio vicino a Yucca Mountain, nel Nevada, sono il residuo di roventi flussi piroclastici prodotti da supereruzioni di circa 12,8 (strato più basso) e 12,7 milioni di anni fa (strato più alto).

Ilya N. Bindeman

QUESTA GRANDE PARETE DI ROCCIA grigia nel Nebraska occidentale si è formata da una pila di ceneri lasciate da una supereruzione avvenuta in un sito ignoto 28 milioni di anni fa.

Rick Otto/University of Nebraska State Museum

SCHELETRI DI ANIMALI RIMASTI SEPOLTI NELLE CENERI DI UN'ERUZIONE catastrofica avvenuta 12 milioni di anni fa nell'Idaho affiorano nel Nebraska nord-orientale, nello Ashfall Fossil Beds State Historical Park. Con ogni probabilità la maggior parte degli animali morì lentamente via via che la cenere, ricadendo al suolo, riempiva i loro polmoni e consumava le loro dentature. Inoltre i composti chimici tossici presenti nella cenere potrebbero avere avvelenato i bacini d'acqua dove gli animali si dissetavano.

to che la cenere galleggia intaserebbe le principali vie d'acqua. Il trasporto lungo i grandi fiumi potrebbe esserne completamente impedito, e in effetti una recente perforazione petrolifera nel Golfo del Messico si è imbattuta in uno spesso strato di detriti supervulcanici presso il delta del Mississippi: a più di 1000 miglia dalla loro origine, Yellowstone. Solo galleggiando sulle acque del fiume per poi aderire ai sedimenti che si sono posati sul fondo oceanico si sono potuti accumulare così tanti detriti a una tale distanza dal vulcano che li aveva eruttati.

Il duro ritorno alla normalità

I ricercatori hanno ragione di credere che possano insorgere (e persistere per molti anni) anche altre gravi conseguenze, provocate dai grandi volumi di gas immessi nell'atmosfera superiore. Ma nuovi studi indicano che alcuni di questi esiti potrebbero anche non essere disastrosi come previsto dai vecchi modelli, anche se altri potrebbero essere peggiori del temuto. Ancora una volta, lo studio della composizione di piccoli sottoprodotti delle eruzioni passate è stato illuminante.

Dei vari gas che costituiscono le eruzioni vulcaniche, il biossido di zolfo (SO_2) è quello che causa l'effetto più intenso sull'ambiente: reagisce infatti con ossigeno e acqua per produrre minuscole goccioline di acido solforico (H_2SO_4), che rappresentano il principale schermo alla radiazione solare, e sulla scia di una supereruzione potrebbero indurre un brusco raffreddamento del clima. Sapendo che il ciclo idrologico impiega mesi o addirittura anni per eliminare completamente le goccioline di acido, molti ricercatori hanno elaborato previsioni apocalittiche di «inverni vulcanici» della durata di decenni, se non di secoli. Ma negli anni recenti altri ricercatori hanno acquisito elementi a favore di una drastica riduzione della durata di questi eventi.

Quasi sempre, tracce dell'acido solforico prodotto dopo grandi eruzioni vulcaniche rimangono intrappolate nella neve e nel ghiaccio. Nel 1996 alcuni ricercatori, studiando carote di ghiaccio provenienti dalla Groenlandia e dall'Antartide, individuarono il picco di acido solforico successivo alla supereruzione di Toba, avvenuta 74.000 anni fa. Quell'eruzione espulse 2800 chilometri cubi di lava e ceneri e ridusse le temperature medie globali tra 5 e 15 gradi. Le conseguenze di questa ondata di freddo furono gravi, ma non durevoli come si era ipotizzato in passato; l'acido solforico presente nel ghiaccio è scomparso dopo solo sei anni, ma alcuni ricercatori ritengono perfino che la situazione si sia risolta nel giro di un anno o due. La buona notizia, insomma, è che probabilmente gli inverni vulcanici sarebbero più brevi di quanto ipotizzato.

E adesso la cattiva notizia. Un nuovo metodo sviluppato per studiare la composizione isotopica dell'ossigeno nelle piogge acide vulcaniche sta rivelando qualcosa di totalmente diverso e allarmante sugli effetti a lungo temine del biossido di zolfo nell'atmosfera. Per trasformarsi in H_2SO_4, occorre che SO_2 si ossidi, cioè che acquisisca due atomi di ossigeno da qualche altro composto già esistente nell'atmosfera. Quale sia esattamente il composto chiamato in causa è tema di un vivace dibattito. Nel 2003 ho iniziato a lavorare con John M. Eiler al California Institute of Technology, e siamo subito andati in cerca di prove da scovare nei miei campioni di ceneri vulcaniche delle eruzioni di Yellowstone e di Long Valley.

Abbiamo iniziato le analisi occupandoci soprattutto di un ossidante particolarmente efficiente: l'ozono, una molecola gassosa costituita da tre atomi di ossigeno, nota per la sua capacità di schermare la Terra dai pericolosi raggi ultravioletti del Sole. Per effetto di alcune trasformazioni chimiche che certi gas subiscono in presenza di quell'intensa radiazione solare, l'ozono è caratterizzato da un'anomalia, che in parole povere può essere descritta come un eccesso di ossigeno-17 nella molecola.

Quando l'ozono o una qualsiasi altra molecola ricca di ossigeno presente nella stratosfera interagisce con SO_2, trasferisce il suo rapporto isotopico dell'ossigeno all'acido risultante: vale a dire l'anomalia dell'ossigeno-17 persiste nel nuovo acido. Nel 2003 i geochimici dell'Università della California a San Diego hanno trovato la prima prova del fatto che questo rapporto isotopico sia anche mantenuto negli atomi di ossigeno dell'acido che più tardi ricade come pioggia e nei composti solfatici che si formano quando la pioggia acida reagisce con la cenere già presente al suolo.

L'eccesso di ossigeno-17 e altre caratteristiche geochimiche che abbiamo riscontrato nei solfati dei campioni di cenere di Yellowstone e della Long Valley implicavano che quantità significative di ozono stratosferico erano state assorbite nelle reazioni con i gas sprigionati dalle supereruzioni in quelle regioni. Altri studi sulle ceneri di eruzioni giganti più antiche nel Colorado e nel Nebraska hanno dimostrato questo fenomeno. Quindi sembra che le emissioni dei supervulcani continuino a causare buchi nello strato di ozono per periodi anche più lunghi di quelli per cui si protrae la loro influenza sul clima.

Questa perdita dell'ozono protettivo farebbe presumibilmente aumentare la quantità di radiazioni ultraviolette dannose in arrivo al suolo, e di conseguenza i danni genetici causati dalle radiazioni stesse. L'entità e la durata della potenziale distruzione dell'ozono sono ancora oggetto di dibattito. Le informazioni raccolte dalle carote di ghiaccio hanno rivelato una diminuzione dal tre all'otto per cento dello strato di ozono in seguito all'eruzione del 1990 del vulcano Pinatubo. Ma che cosa accadrebbe dopo un evento 100 volte più grande? La semplice aritmetica non basta a risolvere il problema, perché i dettagli dell'ossidazione atmosferica sono estremamente complessi e poco chiari.

Le tecniche scientifiche per lo studio e il monitoraggio dei vulcani di tutte le dimensioni si stanno sviluppando con tutta la rapidità possibile. Ma per quanto si riesca a imparare non possiamo impedire un'eruzione. E tutto ciò che si può dire sulle conseguenze degli eventi più catastrofici è pura speculazione. La notizia buona è che ora i ricercatori ne sanno abbastanza sui siti di possibili eruzioni da poter affermare con ragionevole sicurezza che una catastrofe di quelle dimensioni è molto lontana.

DISTRUZIONE DELL'OZONO

NASA Johnson Space Center (spettro atmosferico); Ilya N. Bindeman (fonte grafico)

I pericolosi gas emessi nel 1991 dal vulcano Pinatubo, nelle Filippine, appaiono colorati nelle immagini da satellite dell'atmosfera superiore (*sfondo*). Nuovi studi suggeriscono che gas simili emessi in future eruzioni di supervulcani potrebbero danneggiare notevolmente lo strato di ozono che protegge la Terra, prima di precipitare al suolo come piogge acide e mescolarsi con ceneri a formare solfato. I campioni di solfato estratti da quattro depositi supervulcanici presentano un insolito eccesso di ossigeno-17 (le aree colorate irregolarmente nel grafico rappresentano raccolte di misure); questa abbondanza si riscontra solo in composti che hanno acquisito gli atomi rari nel corso di reazioni con particolari gas presenti nell'atmosfera superiore, con ogni probabilità ozono. I materiali che hanno origine al suolo e vi rimangono, come i prodotti della maggior parte delle piccole eruzioni, non mostrano questa anomalia (*linea blu*).

L'AUTORE

ILYA N. BINDEMAN è geochimico e professore aggiunto al Dipartimento di scienze geologiche dell'Università dell'Oregon. Dopo aver conseguito il dottorato di ricerca all'Università di Chicago nel 1998, ha iniziato a studiare i microscopici cristalli contenuti nelle ceneri vulcaniche alla ricerca di indizi sull'origine e gli effetti delle più imponenti eruzioni del pianeta.

PER APPROFONDIRE

BINDEMAN I.N. e VALLEY J.W., *Low-$\delta^{18}O$ Rhyolites from Yellowstone: Magmatic Evolution Based on Analyses of Zircon and Individual Phenocrysts*, in «Journal of Petrology», Vol. 42, pp. 1491-1517, 2001.

HUIMING B., THIEMENS M.H., LOOPE D.B. e YUAN X.L., *Sulfate Oxygen-17 Anomaly in an Oligocene Ash Bed in Mid-North America: Was It the Dry Fogs?*, in «Geophysical Research Letters», Vol. 30, pp. 1843-1848, 2003.

BINDEMAN I.N., EILER J.M., WING B. e FARQUHAR J., *Sulfur and Triple-Oxygen Isotope Geochemistry of Volcanogenic Sulfate Aerosols*, in «Earth and Planetary Science Letters» (in preparazione, 2006).

I telescopi del futuro

Gli strumenti di domani avranno specchi da 20, 30 e persino 100 metri. La speranza è di costruirli entro dieci anni: perché in astronomia le dimensioni contano

di Roberto Gilmozzi

Very Large Telescope (senza ottica adattiva)
Dimensione delle specchio: 8,2 metri
Risoluzione: 0,4 secondi d'arco
Tempo di esposizione: 620 secondi

Hubble Space Telescope
Dimensione delle specchio: 2,4 metri
Risoluzione: 0,04 secondi d'arco
Tempo di esposizione: 1600 secondi

Very Large Telescope (con ottica adattiva)
Dimensione delle specchio: 8,2 metri
Risoluzione: 0,012 secondi d'arco
Tempo di esposizione: 160 secondi

OWL Telescope (proposto)
Dimensione delle specchio: 100 metri
Risoluzione: 0,001 secondi d'arco
Tempo di esposizione: 1 secondo

Alcuni dei momenti migliori che trascorro all'osservatorio del Cerro Paranal, in Cile, sono di notte, quando, dopo un giorno di lavoro, vado sul «ponte», il nome con cui chiamiamo la piattaforma che ospita i quattro telescopi da otto metri di diametro del Very Large Telescope (VLT). È un momento magico: la vastità del cielo stellato, i lenti movimenti delle cupole, il piacere di fumare la pipa, il deserto, scuro e appena visibile, proiettato sull'orizzonte. Mentre ammiro il VLT, il gruppo di telescopi più avanzato del mondo, con le sue quattro macchine da 430 tonnellate che ruotano in silenzio in un complicato balletto con il cielo, rifletto su quanto sono fortunato a essere coinvolto in un progetto così grandioso. È un successo dell'umanità. Come gli altri grandi telescopi di oggi, il Keck, lo Hubble Space Telescope e il Very Large Array, il VLT incarna le tecnologie più avanzate che la nostra civiltà può offrire.

DA UNA LUMINOSITÀ INDISTINTA A UN TRIONFO DI STELLE. Un telescopio da 100 metri di diametro con ottica adattiva ci permetterà una visione ancora più nitida di quella di Hubble.

Slim Films (*composizione*); European Southern Observatory (*galassia NGC 3603*)

QUATTRO SECOLI A OSSERVARE IL CIELO

Telescopi come OWL e TMT proseguirebbero la lunga tendenza storica all'aumento delle dimensioni dell'apertura.

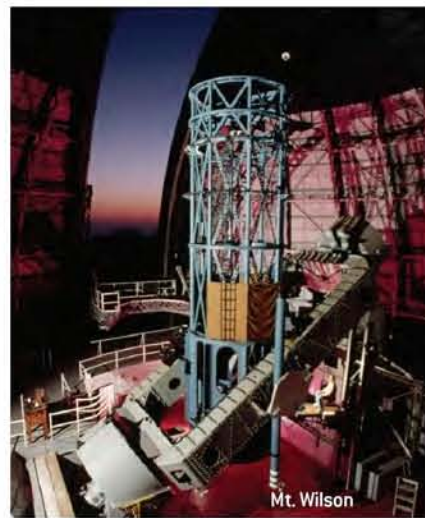
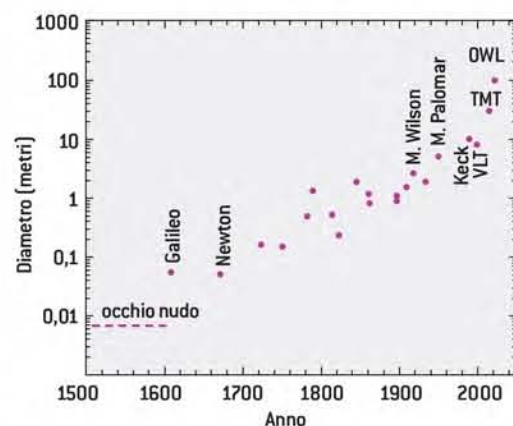
Melissa Thomas (grafico); Corbis (telescopi)

Galileo

Newton

Mt. Wilson

Mt. Palomar

Keck 1

L'AUTORE

ROBERTO GILMOZZI è responsabile scientifico dello studio di progetto di OWL. Dal 1999 al 2005 è stato anche direttore del Very Large Telescope Observatory di Cerro Paranal, in Cile, dello European Southern Observatory. I suoi interessi scientifici comprendono le esplosioni di supernova e i loro resti, il fondo cosmico nei raggi X, la storia della formazione delle stelle e dell'universo.

Ma gli astronomi non si fermano mai. Il VLT non era ancora stato terminato che molti di noi iniziarono a pensare ai suoi successori, telescopi i cui specchi principali misurerebbero 25, 30 o anche 100 metri di diametro. Un'idea in cui sono stato profondamente coinvolto è un mostro chiamato OWL (per la sua acuta visione notturna, visto che significa «gufo» in inglese, e perché sarà OverWhelmingly Large, «esageratamente grande») che riempirebbe quasi l'intero ponte del Paranal con il suo specchio di 100 metri di diametro.

Come tutti i nuovi strumenti scientifici, i telescopi da 8-10 metri attuali non stanno solo rispondendo agli interrogativi per cui erano stati costruiti, ma ne stanno ponendo di nuovi, più profondi e stimolanti, che richiedono strumenti ancora più grandi. Analizzare la composizione dei pianeti di tipo terrestre in altri sistemi solari e cercare tracce di vita; studiare le prime galassie che si formarono nell'universo; capire la natura della materia e dell'energia oscure; riprendere immagini della moltitudine di oggetti del nostro sistema solare che non sono stati avvicinati dalle sonde spaziali. Tutto ciò spinge gli astronomi verso una generazione di telescopi ottici giganti con capacità centinaia o migliaia di volte superiori a quelle disponibili oggi. Molte istituzioni europee hanno riconosciuto la priorità di costruire un simile telescopio, mentre per l'Accademia delle scienze statunitense è secondo solo al successore di Hubble, il James Webb Space Telescope (JWST). Al momento sono in fase di studio diversi progetti, compresi OWL, il Thirty Meter Telescope (TMT, di 30 metri) e il Giant Magellan Telescope (GMT), di 24 metri.

Storicamente, i telescopi hanno seguito una loro versione della legge di Moore: ogni generazione è grande circa il doppio della precedente, e richiede vari decenni per essere realizzata. Questo andamento è illustrato particolarmente bene dalla cosiddetta «California progression», una crescita delle dimensioni verificatasi durante il XX secolo: il telescopio Hooker da 2,5 metri di Monte Wilson (1917), il telescopio Hale da cinque metri di Monte Palomar (1948) e i telescopi gemelli Keck da dieci metri di Mauna Kea, alle Hawaii (1992). Seguendo questo precedente, i telescopi della prossima generazione dovrebbero avere un diametro di 20 metri e diventare operativi attorno al 2025.

Spazio di miglioramento

Ma allora chi propone telescopi da 30 o anche 100 metri per la metà del prossimo decennio è da prendere per pazzo? Uno sguardo più approfondito alla sfida che si dovrà affrontare per costruire un simile telescopio potrebbe rassicurarci riguardo alla salute mentale degli astronomi. Non solo l'esigenza di telescopi terrestri più grandi è pressante ma, soprattutto, la tecnologia necessaria è già quasi tutta disponibile.

Un forte stimolo a infrangere la legge di Moore dei telescopi è che gli astronomi stanno esaurendo le altre opzioni per migliorare la capacità di raccogliere luce del loro strumenti. In un telescopio riflettore, la luce rimbalza sullo specchio primario e poi colpisce il secondario, che la focalizza in un punto conveniente, dove si può guardarla con gli occhi, dividerla in un arcobaleno di colori per l'analisi spettroscopica o scattare una fotografia. Quando gli astronomi parlano delle dimensioni di un telescopio si riferiscono al diametro dello specchio primario. Raddoppiarlo permette di vedere oggetti quattro volte più deboli, oppure un corpo di una data luminosità a una distanza doppia.

Negli ultimi cinquant'anni i telescopi sono diventati più sensibili agli oggetti deboli, non solo perché è aumentato il diametro degli specchi, ma anche grazie ai progressi dei sensori. Quando fu costruito, il telescopio Hale da cinque metri era equipaggiato con lastre fotografiche, che registrano solo una piccola percentuale della luce che le colpisce. I sensori elettronici di oggi hanno un'efficienza quasi del 100 cento: un miglioramento della sensibilità pari a un aumento del diametro di un fattore cinque. In effetti, è come se i telescopi di oggi fossero dieci volte più grandi dei loro predecessori. Ma perché la prossima generazione compia lo stesso balzo, ora che restano pochi margini per migliorare l'efficienza dei sensori, il diametro deve raggiungere i 100 metri.

Il dibattito su quale sia il diametro più grande raggiungibile è acceso, ma nessuno mette in dubbio la necessità di spingere ulteriormente in avanti le dimensioni della prossima generazione. Tradizionalmente, le dimensioni dei nuovi telescopi sono state limitate dalla capacità di produrre lo specchio di vetro, dargli la giusta forma e smerigliarlo. La luce visibile ha una lunghezza d'onda inferiore rispetto alle onde radio, e per questo motivo anche se le antenne radio possono essere enormi i loro requisiti sono meno stringenti rispetto a quelli degli specchi ottici, così come c'è bisogno di un controllo motorio più preciso per tenere un granello di sabbia tra le dita che per trasportare un masso.

Il telescopio Hale da cinque metri ha uno specchio parabolico la cui superficie è stata lavorata con una precisione di 50 nanometri. Se fosse grande quanto l'Oceano Atlantico, l'irregolarità più grande della sua superficie sarebbe alta cinque centimetri. Per ridurne le irregolarità della superficie, i costruttori usarono uno strumento coperto di pece e, nelle fasi finali, alcune zone furono rifinite a mano. L'impresa richiese 11 anni, durante i quali la superficie era controllata ogni due giorni.

Gli specchi di oggi sono modellati con il controllo di un computer, il che accelera enormemente i tempi. I quattro specchi da 8,2 metri del VLT sono stati lucidati in meno di un anno, mentre la loro superficie era controllata quasi in continuazione. La qualità della loro superficie equivale, o supera di poco, quella del telescopio Hale, anche se la forma (un iperboloide, che produce il fuoco più nitido possibile) è molto più complessa. La lucidatura, quindi, non è più l'ostacolo maggiore.

Il problema più serio è la fabbricazione del vetro. Per realizzare vetri di otto metri di diametro i produttori di telescopi hanno dovuto costruire apposite fabbriche, e spesso buttare via diversi specchi prima di ottenerne uno buono. Le procedure di oggi non basterebbero nemmeno per raddoppiare le dimensioni.

Per fortuna l'astronomo italiano Guido Horn D'Arturo ha trovato la soluzione già dal 1932: uno specchio a segmenti. Gli specchi dei telescopi gemelli Keck, per esempio, sono un mosaico di 36 segmenti, ciascuno dei quali è un esagono del diametro di 1,8 metri. La forma esagonale permette di incastrarli creando la superficie di un iperboloide: ciascun segmento ha un profilo leggermente differente dagli altri, che dipende dalla sua distanza dal centro dello specchio. In teoria, un simile progetto può essere scalato a qualsiasi dimensione, ma ha il problema di allineare tutti i pezzi con una precisione migliore della lunghezza d'onda, per minimizzare gli effetti delle connessioni fra i segmenti sulla qualità dell'immagine e per tenerli insieme nonostante il vento.

Come il Keck, OWL e TMT sarebbero costruiti con segmenti esagonali, mentre i progettisti di GMT hanno scelto una strada diversa: per minimizzare i problemi imposti dalla segmentazione, hanno scelto di costruire meno segmenti ma più grandi. Il loro telescopio sarebbe un mosaico di sette elementi circolari di specchio del diametro di 8,4 metri (il primo dei quali è già in costruzione, come dimostrazione di principio). Il problema di questo metodo è che è difficile da scalare ulteriormente.

C'è l'esigenza di **telescopi terrestri più grandi,** e la tecnologia è già disponibile

In sintesi/*Very Very Large Telescopes*

- In astronomia, le dimensioni sono importanti. I grandi telescopi permettono di vedere oggetti più deboli e producono immagini più dettagliate. I telescopi più grandi di oggi per la luce visibile e il vicino infrarosso hanno specchi di diametro compreso fra 8 e 10 metri, e i ricercatori stanno lavorando alla prossima generazione, da 20 metri fino a un gigantesco specchio da 100 metri di diametro.
- Situati a terra, i nuovi telescopi sarebbero equipaggiati con ottiche adattive, per annullare gli effetti della turbolenza atmosferica. Produrrebbero immagini più nitide di quelle di Hubble, e a costi decisamente più bassi. I nuovi strumenti potrebbero effettuare ricerche ben al di là delle capacità dei telescopi di oggi, come cercare pianeti di tipo terrestre e analizzare la composizione di quelli eventualmente scoperti.

In cerca della visione

Un'altra caratteristica importante di un telescopio è la risoluzione, ovvero la capacità di distinguere i dettagli fini. In linea teorica, un grande telescopio dovrebbe avere entrambe le qualità. Più il telescopio è grande e meno le sue immagini sono degradate dalla diffrazione, un effetto che si verifica quando le onde in arrivo sono tagliate al bordo esterno dello specchio. Fino a poco tempo fa però la diffrazione è stata un elemento controverso per i telescopi ottici terrestri. Anche nei siti migliori, la turbolenza dell'aria sfuoca tutte le strutture più piccole di 0,3 secondi d'arco. Se guardaste la stella gigante Betelgeuse (che ha un diametro di 0,05 secondi d'arco) attraverso il telescopio da cinque metri di monte Palomar, tutto ciò che vedreste è un puntino scintillante di luce rossa: che è più brillante, ma non più chiaro di ciò che vedreste con un telescopio amatoriale da 20 centimetri o persino a occhio nudo.

I telescopi in orbita hanno il problema opposto, perché producono immagini con una risoluzione spettacolare, ma non hanno la sensibilità per vedere gli oggetti più deboli. Il diametro di Hubble fu limitato a 2,4 metri dalle dimensioni del vano di carico dello space shuttle e il JWST avrà uno specchio di appena 6,5 metri.

Questo compromesso tra sensibilità e risoluzione non è abbastanza per la prossima generazione di telescopi, per i cui obiettivi scientifici servono entrambi. In un'esposizione notturna, un telescopio da 100 metri potrebbe vedere oggetti celesti con una luminosità pari a un millesimo dei limiti attuali. Dove i telescopi di oggi vedono un angolo oscuro dello spazio, vedrebbe un affollamento di oggetti deboli, ma, senza un'alta risoluzione, questi oggetti inizierebbero a confondersi tra loro.

Una combinazione di risoluzione e sensibilità è importante anche per l'analisi dei pianeti extrasolari di tipo terrestre. Per vedere un simile pianeta, con circa un miliardesimo della luminosità della sua stella, gli astronomi dovrebbero oscurare la stella con un disco opaco, il «coronografo». Se il disco fosse troppo grande, però, nasconderebbe anche il pianeta. Un'alta risoluzione significa che gli astronomi possono usare un disco più piccolo, estendendo il raggio della loro ricerca di pianeti a distanze inferiori dalla stella.

Le dimensioni minime di un telescopio che può scrutare i nostri dintorni galattici alla ricerca di pianeti in orbite simili a quelle della Terra sono di circa 80 metri. Potrebbe scandagliare un volume che comprende circa 400 sistemi con stelle simili al Sole e studiare lo spettro dei pianeti di tipo terrestre, se ci fossero, di circa 40 di esse. Uno strumento di 30 metri potrebbe studiare poche dozzine di sistemi e per analizzare lo spettro di uno di questi dovrebbe raccogliere luce per settimane.

Evolvere e adattarsi

Per raggiungere una risoluzione così alta, il telescopio dovrà affidarsi alle ottiche adattive per eliminare le distorsioni introdotte dalla turbolenza atmosferica. L'idea è di tenere sotto controllo una stella di riferimento e mantenerla perfettamente a fuoco regolando la superficie di uno specchio. Questo specchio può essere quello secondario, oppure un altro specchio più piccolo interposto fra il secondario e i rivelatori. Una serie di piccoli pistoni, o attuatori, agisce sul retro dello specchio per regolarne accuratamente la forma.

Questo sistema permette ai telescopi di lavorare al massimo o vicino al massimo del loro potere risolutivo, limitati solo dalla diffrazione, come se l'atmosfera non ci fosse. Un telescopio di 100 metri dovrebbe essere capace di risolvere strutture di 0,001 secondi d'arco, circa 40 volte meglio di quanto può fare Hubble. Osservata con un simile strumento, Betelgeuse non apparirebbe più come un semplice punto di luce, ma come un'immagine di 3000 pixel con un livello di dettaglio oggi disponibile solo per i pianeti vicini.

La tecnica è già in uso in molti telescopi, ma per farla funzionare su strumenti ancora più grandi è necessario scalarla. E non è detto che sia possibile: un sistema di ottiche adattive su un telescopio di 100 metri richiederebbe 100.000 attuatori, mentre i sistemi di oggi ne hanno al massimo 1000. Il computer di controllo dovrebbe riuscire ad aggiornare la forma di questo specchio centinaia di volte al secondo, e per ora la tecnologia dei processori non è all'altezza del compito.

I tecnici stanno adottando un approccio per fasi, costruendo prima sistemi di ottiche adattive per lavorare nell'infrarosso, cosa che richiede meno attuatori per-

RADIOASTRONOMIA DA RECORD

Mentre sono in fase di progettazione nuovi telescopi ottici sempre più grandi, che apriranno nuove finestre nelle lunghezze d'onda del visibile, il mondo della radioastronomia si prepara a una vera a propria rivoluzione. Nel 2020, quando lo Square Kilometer Array (SKA) sarà completato, sarà il più grande telescopio mai costruito. Con la sua area equivalente di un milione di metri quadrati (ovvero un chilometro quadrato, da cui il nome) sarà 100 volte più sensibile di ogni altro radiotelescopio e 10.000 volte più veloce.

SKA Team/W.J. Kleppe

SKA Team/ASTRON

ALLE FRONTIERE DELLA TECNOLOGIA. Al centro, raffigurazione artistica di una possibile configurazione per gli elementi centrali di SKA. L'Aperture Plane Array sarà in grado di seguire contemporaneamente molteplici oggetti con differente risoluzione, guidata elettronicamente da sofisticati programmi informatici in grado di combinare il segnale proveniente da ogni elemento. Nel disegno in basso, ecco come apparirà la schiera delle centinaia di stazioni di SKA. Nella pagina a fronte, il sito di Guiyang, nella provincia cinese di Guizhou, è una delle quattro località candidate a ospitare il radiotelescopio; la decisione finale sarà presa entro il 2008, la costruzione inizierà nel 2014 e lo strumento dovrebbe essere completato entro il 2020.

Grazie alle sue caratteristiche innovative, SKA avrà allo stesso tempo una grande sensibilità e una squisita risoluzione angolare (oltre un secondo d'arco), pur mantenendo un campo di visione molto grande (quattro volte la superficie della Luna piena ad alte frequenze, oltre 200 volte la superficie della Luna per frequenze al di sotto di un gigahertz).

SKA sarà uno strumento ad apertura di sintesi, in grado cioè di combinare digitalmente il segnale di un grande numero di elementi per simulare un telescopio con un diametro equivalente alla separazione massima fra gli elementi, ovvero 3000 chilometri. I vari elementi saranno distribuiti in alcune centinaia di stazioni, e circa la metà di essi sarà concentrata in un nucleo centrale di cinque chilometri di diametro, mentre gli elementi rimanenti verranno disposti a spirale a distanza crescente dal centro. Le soluzioni tecnologiche più adatte per le differenti stazioni sono in corso di sviluppo, ma il progetto prevede che il nucleo centrale sarà costituito da elementi planari fissi, chiamati *Aperture Plane Array*. È un concetto innovativo, che permetterà di osservare tutta la volta celeste allo stesso tempo e di studiare simultaneamente differenti oggetti grazie a tecniche di interferometria. Le stazioni esterne saranno invece antenne paraboliche orientabili, ciascuna con un diametro tra i 12 e i 15 metri, per un totale di almeno 4000 antenne.

Una parte fondamentale della tecnologia necessaria al funzionamento di SKA sarà la trasmissione digitale di segnali lungo le migliaia di chilometri tra le varie stazioni, e la potenza di calcolo e di memoria necessarie a elaborare i dati. Con un volume di dati stimato in 2 terabyte al secondo sarà necessario sviluppare soluzioni ad alta tecnologia in grado di gestire un flusso di informazione superiore al traffico Internet di tutta l'Europa! In effetti, SKA sarà il primo telescopio digitale della storia, che userà supercomputer oltre 1000 volte più veloci di quelli attuali per combinare ed elaborare i dati provenienti dalle migliaia di elementi sparsi su scala continentale.

Date le sue caratteristiche, SKA nasce come un progetto globale, che vede coinvolti istituti, scienziati, ingegneri e partner industriali provenienti da 18 paesi, fra cui Italia, Stati Uniti, Gran Bretagna, Russia, Australia e Cina. Dopo oltre dieci anni di studi preliminari, il progetto sta entrando nella prima fase operativa. La scelta finale del sito sarà fatta entro il 2008, selezionandolo fra i quattro candidati in lizza: Argentina, Australia, Cina e Sudafrica. Vista la scala del progetto, saranno tenuti in considerazione molti criteri, dal livello di rumore radio del sito alla disponibilità di risorse tecniche e scientifiche nella regione, fino a coinvolgere i più alti livelli dei governi interessati. Grazie a un investimento europeo di 38 milioni di euro, lo sviluppo di un prototipo con un'area dell'1 per cento rispetto alle specifiche di SKA sta avanzando speditamente, e sarà terminato entro il 2010. Questo prototipo fornirà un prezioso banco di prova per le tecnologie e gli strumenti che costituiranno il telescopio finale. Si prevede poi di procedere entro il 2014 alla costruzione del primo dieci per cento di SKA, mentre l'intero progetto dovrebbe essere completato entro il 2020.

Tutto ciò permetterà di osservare l'universo a lunghezze d'onda comprese tra un centimetro e tre metri, con una risoluzione spaziale e temporale impressionante, facendo di SKA uno strumento unico nella sua versatilità e potenza. Grazie a SKA, i cosmologi misureranno la distribuzione dell'idrogeno neutro nell'universo primordiale a oltre 12 miliardi di anni luce di distanza, osservando la nascita delle prime stelle che misero fine all'epoca oscura del cosmo. In pochi mesi di operazione, SKA determinerà la posizione e le proprietà di oltre un miliardo di galassie, fornendo precise indicazioni sulla natura dell'energia oscura e sulla massa dei neutrini. Lo studio dei campi magnetici primordiali permetterà di comprenderne le origini e di verificare l'influenza del magnetismo cosmico sulla formazione delle galassie.

SKA darà anche nuovo impulso alla verifica della teoria della relatività generale, perché sarà in grado di scoprire decine di migliaia di pulsar. Le pulsar sono stelle di neutroni che ruotano molto velocemente su se stesse emettendo lampi radio a intervalli estremamente regolari, e che gli astronomi possono sfruttare come radiofari naturali di grande precisione. Fra esse, le più interessanti saranno quelle in orbita attorno a buchi neri, o attorno al buco nero centrale supermassiccio della nostra galassia, che forniranno un banco di prova formidabile per la teoria della gravità di Einstein. Nel campo dell'astrobiologia, SKA studierà i processi di formazione planetaria in corso attorno ad altre stelle, cercherà pianeti di tipo terrestre fino a una distanza di 150 parsec dalla Terra, a caccia delle emissioni caratteristiche di molecole complesse e amminoacidi necessari per la vita. Inoltre, gli astronomi potranno di captare segnali radio di eventuali civiltà tecnologicamente avanzate provenienti da un milione di stelle di tipo solare nella nostra galassia, con una sensibilità tale da potersi sintonizzare sui programmi tv di ipotetiche civiltà aliene. Senza dimenticare che queste nuove osservazioni apriranno nuove prospettive sull'universo, dandoci risposte a domande che oggi non siamo nemmeno in grado di porre.

Il costo stimato è di circa un miliardo di euro, una somma ingente, all'apparenza, ma un investimento modesto rispetto alle potenziali scoperte che SKA potrà fare. È chiaro che sarà necessario sviluppare nuovi metodi industriali per produrre su vasta scala e a basso costo le migliaia di elementi di SKA, sia a livello dei singoli telescopi che del sistema informatico. Tutto ciò richiede una forte sinergia fra scienziati, tecnici, ingegneri e partner industriali, e diversi gruppi di lavoro sono già all'opera in questo senso. Grazie alla sua natura internazionale e al coinvolgimento di uomini e donne dal mondo intero in questa affascinante sfida, oltre a esplorare l'universo SKA potrà forse contribuire alla pace e alla collaborazione scientifica nel mondo.

ROBERTO TROTTA

ché a grandi lunghezze d'onda (come nel caso dell'infrarosso) l'effetto della turbolenza è minore. Gli astronomi dovrebbero anche riuscire ad approfittare degli sforzi compiuti per sviluppare sistemi avanzati di ottiche adattive per la medicina, la sorveglianza militare e l'elettronica di consumo. Una nuova tecnica promettente è quella delle ottiche adattive multiconiugate, che correggono la turbolenza su un vasto campo visivo, in modo che i telescopi non siano limitati a un piccolo campo di spazio attorno alla stella di riferimento.

L'interferometria, una tecnica che combina la luce di più di un telescopio, può raggiungere una risoluzione ancora più elevata dei grandi telescopi proposti, e un sistema di questo tipo è già in funzione all'osservatorio di Cerro Paranal. I quattro telescopi del VLT sono posti fino a 130 metri di distanza l'uno dall'altro, quindi unire la loro luce offre la stessa risoluzione di un telescopio di 130 metri di diametro. Ma gli interferometri hanno i loro limiti, perché possono osservare solo un campo visivo molto ridotto: usarli è come guardare attraverso una cannuccia. Inoltre, per la complessità delle loro ottiche, sfruttano solo una piccola frazione della luce che raccolgono, contro il 50 per cento o più dei telescopi standard, e in ogni caso la loro area di raccolta totale è pari alla somma dei telescopi che li compongono. In breve, come i telescopi spaziali, guadagnano risoluzione a scapito della sensibilità, e quindi non possono sostituire uno strumento terrestre gigante.

Dadi e bulloni

Un elefante non è fatto come una formica. Il peso di una creatura aumenta con il cubo delle sue dimensioni, mentre la capacità dello scheletro di sostenere il peso aumenta solo con il quadrato: per questo l'elefante in proporzione ha bisogno di gambe molto più grandi. Ciò che vale per i grandi mammiferi terrestri vale anche per i telescopi, e i progressi nella tecnologia ottica sarebbero inutili se il telescopio non fosse in grado di reggere neppure il proprio peso. Anche se i radioastronomi hanno costruito antenne orientabili con diametri fino a 100 metri, un telescopio ottico ha esigenze meccaniche maggiori, perché opera a lunghezze d'onda molto inferiori.

La struttura del telescopio deve essere abbastanza rigida da mantenere gli specchi perfettamente allineati gli uni con gli altri e da resistere alle vibrazioni prodotte dal vento. In generale, i telescopi corti e larghi tendono a essere più rigidi di quelli lunghi e stretti, ma esigono che la luce sia deviata di più per essere portata a un fuoco, e questo complica la progettazione dell'apparato ottico. Così, va trovato un equilibrio tra la frequenza di risonanza e la lunghezza focale. Il VLT oscilla un po' al passaggio del vento, ma lo specchio secondario cancella gli effetti di queste vibrazioni muovendosi nella direzione opposta fino a 70 volte al secondo. Il progetto di OWL prevede un meccanismo simile.

Un altro problema è che mentre il telescopio perlustra il cielo il suo peso si sposta, e questo può far flettere lo strumento e disallineare gli specchi. Molti grandi telescopi di oggi sono costruiti con una struttura a scheletro progettata negli anni trenta per Palomar. In questa struttura, ogni specchio è sorretto da un telaio aperto, simile a una scatola, costituito da quattro travi triangolari. Quando sono inclinati, i telai flettono e gli specchi si spostano lateralmente, ma poiché tutti gli specchi sono sorretti dallo stesso tipo di telaio si spostano della stessa misura, rimanendo allineati. Il progetto di OWL adotta un approccio simile, ma ha il vantaggio di poter essere costruito da pochi componenti standard.

Il peso totale della struttura sarà tra 10.000 e 30.000 tonnellate: per fare un confronto, la Torre Eiffel pesava circa 10.000 tonnellate, quando fu costruita. Ma per quanto enorme possa sembrare OWL sarebbe molto più leggero dei telescopi attuali, in proporzione. Se uno dei telescopi del VLT fosse scalato fino alle dimensioni di OWL peserebbe 500.000 tonnellate. In ogni caso, muovere 10.000 tonnellate con la precisione necessaria per un telescopio è una sfida. Le opzioni comprendono carrelli simili a quelli dei treni con un movimento a frizione su ciascuna ruota, sottili strati di olio su cui il telescopio galleggerebbe e la levitazione magnetica.

Tecnicamente, insomma, gli astronomi non sono pazzi a voler costruire telescopi con diametri fino a 100 metri. Il problema principale è il costo. Storicamente, il costo di un telescopio è proporzionale al diametro del suo specchio principale elevato alla 2,6 ($D^{2,6}$). Quindi se i quattro telescopi da otto metri del VLT sono costati circa 100 milioni di dollari ciascuno, un telescopio di 20 metri costerebbe circa un miliardo di dollari e uno da 100 metri l'incredibile somma di 70 miliardi di dollari. Se la legge dei costi rimanesse valida, allora gli astronomi dovrebbero prendere in seria considerazione l'idea di costruire molte copie di un telescopio più piccolo, fino a raggiungere dimensioni equivalenti: allora il costo crescerebbe come D^2. Per un miliardo di dollari potremmo costruire dieci telescopi da 8,2 metri, con un'area equivalente a un singolo telescopio da 26 metri. Purtroppo, per le ragioni già illustrate, dimensioni uguali non significa uguale capacità. Se fossero usati come singoli strumenti, l'insieme dei telescopi avrebbe la sensibilità di un telescopio da 26 metri, ma la risoluzione di uno da 8,2. Usati come interferometro, avrebbero una risoluzione più elevata, ma una sensibilità molto inferiore.

Il problema principale nel costruire un telescopio gigantesco **non è tecnico ma economico**

GLI ALTRI PROGETTI

THIRTY-METER TELESCOPE (TMT)

Diametro: 30 metri
Costo stimato: 700 milioni di dollari
Progetto: Primario segmentato iperboloidale
Sito Web: www.tmt.org

LARGE-APERTURE MIRROR ARRAY (LAMA)

42 metri (area), 54 metri (risoluzione)
50 milioni di dollari
18 specchi paraboloidali da 10 metri a mercurio liquido
www.astro.ubc.ca/LMT/lama

GIANT MAGELLAN TELESCOPE (GMT)

21,4 metri (area), 24,5 metri (risoluzione)
500 milioni di dollari
Sette specchi iperboloidali da 8,4 metri
www.gmto.org

EURO50

50 metri
700 milioni di dollari
Primario segmentato ellissoidale
www.astro.lu.se/~torben/euro50

Todd Mason, Mason Productions (TMT); Paul Hickson, University of British Columbia (LAMA); Todd Mason, Mason Productions e Carnegie Observatories (GMT); Euro50 Project, Osservatorio di Lund, Svezia (Euro50)

Un modo per contenere i costi sarebbe produrre i componenti in serie. Questo richiede però un nuovo approccio all'ottica. Invece del solito specchio primario iperbolico, per cui andrebbe costruito appositamente ciascun segmento a seconda della sua posizione, un telescopio da 100 metri potrebbe avere uno specchio sferico, i cui i segmenti hanno tutti la stessa forma. Una catena di montaggio potrebbe produrre i 3048 elementi al ritmo di uno ogni due giorni. Lo svantaggio è che uno specchio sferico introduce una distorsione dell'immagine, ed è necessario dotare il telescopio di un «correttore». Ma anche così il sistema sarebbe molto più economico.

Una delle voci di spesa più importanti di ogni telescopio è la cupola che lo protegge dagli eventi atmosferici. Palomar ha una cupola più grande della basilica di San Pietro. Un motivo per cui è tanto grande è che la montatura è inclinata per puntare verso la stella polare, in modo che il telescopio possa inseguire le stelle semplicemente ruotando attorno al proprio asse. I telescopi moderni sono sorretti da montature più compatte chiamate «altazimutali», che richiedono un meccanismo di controllo più complicato, reso possibile dai computer. Ma anche con il sistema altazimutale, un telescopio di 100 metri richiederebbe una cupola enormemente costosa. Inoltre, una struttura così monumentale potrebbe generare le sue sacche di turbolenza. Così OWL avrà solo un hangar su rotaie per coprire il telescopio durante il giorno o con il brutto tempo. Il telescopio opererà all'aria aperta e potrebbe resistere a un vento moderatamente forte, fino a 15 metri al secondo.

Il costo di un OWL di 100 metri sarebbe di circa 1,2 miliardi di dollari, mentre quello di TMT si stima in circa 700 milioni di dollari e di GMT in 400 milioni. Poiché sono comunque somme ingenti, sarà probabilmente essenziale una collaborazione internazionale.

Strumenti per il XXI secolo

L'astronomia sta vivendo la sua età dell'oro, ma possiamo aspettarci di più, entro il 2015. Sensori innovativi e ottiche adattive estenderanno le capacità dei telescopi da 8-10 metri, e nuovi sensori elettronici e spettrometri hanno iniettato altra vita in Hubble. L'interferometria inseguirà oggetti sempre più deboli e raggiungerà una risoluzione migliore del millesimo di secondo d'arco. E poi sarà lanciato JWST, specializzato in osservazioni nell'infrarosso. L'Atacama Large Millimeter Array, ponte fra la radioastronomia e l'astronomia nell'infrarosso, avrà iniziato a lavorare e proseguirà il progetto dello Square Kilometre Array (*si veda il box a pp. 56-57*).

Con tutti questi progressi, ci sarà davvero bisogno di un nuovo, gigantesco telescopio ottico? La risposta è sì. Problemi cruciali come lo studio dei pianeti extrasolari e dei blocchi di costruzione delle stelle e delle galassie non possono essere risolti in altro modo. Per lo studio nella luce visibile o nel vicino infrarosso, i telescopi terrestri offrirono una risoluzione e una sensibilità più alte e a un costo inferiore rispetto a quelli orbitali. Di recente un comitato internazionale ha esaminato il progetto di OWL e concluso che è fattibile ma rischioso, sia tecnicamente sia economicamente. Ora il gruppo sta lavorando a un progetto più piccolo, e una decisione è attesa per la fine dell'anno. I vari progetti potrebbero convergere: lo stesso TMT è una fusione di progetti precedenti.

Nei secoli, i telescopi sono passati dalle dimensioni di un comodino a quelle di una stanza, di una casa, di una cattedrale e ora di un grattacielo. Oggi possiamo costruire strumenti in grado di vedere le prime stelle nate nell'universo e i pianeti attorno ad altre stelle. Il problema non è più se siamo capaci di costruire telescopi giganti e perché, ma quando farlo e di quale grandezza.

PER APPROFONDIRE

GILMOZZI R. e DIERICKX P., *OWL Concept Study*, in «ESO Messenger», n. 100, pp. 1-10, giugno 2000. Online sul sito www.eso.org/projects/owl/publications/2000-05-Messenger.htm.

DEGRASSE TYSON N., *The Light Brigade*, in «Natural History», Vol. 115, n. 2, pp. 18-29, marzo 2006.

Exploring the Cosmic Frontier: Astrophysical Instruments for the 21st Century, ESO Astrophysics Symposia, Springer-Verlag (in stampa). www.mpifr-bonn.mpg.de/berlin04/.

La scienza del SUDOKU

Per risolvere un sudoku non serve la matematica, e nemmeno l'aritmetica. Eppure questo gioco divenuto così popolare propone molti affascinanti problemi matematici

di Jean-Paul Delahaye

Si potrebbe pensare che un gioco basato sulla logica attiri pochissime persone. Probabilmente matematici, appassionati di computer, amanti del gioco d'azzardo. Eppure in pochissimo tempo il *sudoku* è diventato straordinariamente popolare, tanto da ricordare la moda del cubo di Rubik in voga all'inizio degli anni ottanta.

A differenza del cubo di Rubik, tridimensionale, il sudoku è uno schema quadrato a due dimensioni. In genere è composto da 81 caselle (disposte su nove righe e nove colonne) ed è diviso in nove quadrati più piccoli, ciascuno contenente nove caselle, che chiameremo sottogriglie. All'inizio del gioco in alcune caselle sono stampati dei numeri, il giocatore deve riempire le caselle vuote con le cifre da 1 a 9 in modo che nessuna cifra compaia due volte nella stessa riga, colonna o sottogriglia. Ogni schema ha una soluzione unica.

Paradossalmente, nonostante sia un gioco basato sui numeri, il sudoku non richiede da parte dei giocatori neppure un briciolo di conoscenza matematica. Infatti, non è richiesta alcuna operazione per completare uno schema che, in teoria, potrebbe essere composto con un qualsiasi insieme di nove simboli diversi (lettere, colori, icone e così via). Ciò nonostante il sudoku è, per matematici e informatici, fonte di numerosi e stimolanti problemi.

L'antenato del sudoku non è, come di solito si crede, il quadrato magico, ovvero una matrice in cui la somma dei numeri su ogni riga, su ogni colonna e sulle diagonali è sempre la stessa. Anzi, a parte i numeri e lo schema quadrato, il sudoku non ha quasi nulla a che vedere con il quadrato magico, mentre ha moltissimo a che vedere con il quadrato latino (*si veda il box a p. 62*).

Un quadrato latino di ordine n è una matrice di n^2 caselle (n per lato) riempite con n simboli in modo che lo stesso simbolo non compare mai due volte nella stessa riga o nella stessa colonna (quindi ognuno degli n simboli è usato esattamente n volte). Le origini di questi schemi risalgono al Medioevo, ma fu il matematico svizzero Leonhard Euler a chiamarli quadrati latini e a studiarne le proprietà, nel XVIII secolo.

Un sudoku standard è come un quadrato latino di ordine 9, con unica differenza, cioè il vincolo che in ogni sottogriglia devono essere presenti i numeri da 1 a 9. Il primo gioco di questo genere apparve nel numero del maggio 1979 della rivista «Dell Pencil Puzzles and Word Games», e secondo le ricerche svolte da Will Shortz, curatore della rubrica di cruciverba del «New York Times», pare sia stato creato da un architetto in pensione di nome Howard Garns, morto a Indianapolis nel 1989 (o nel 1981, ci sono diverse versioni): troppo presto per assistere al successo mondiale della sua invenzione.

Nel 1984 il gioco, pubblicato dalla Dell Magazines con il nome di «Number Place», approdò a una rivista giapponese che finì con il chiamarlo sudoku, termine che più o meno significa «singoli

SUDOKU PER VOI. Inserite un numero da 1 a 9 in ogni casella in modo che nessun numero compaia due volte nella stessa riga, colonna o sottogriglia (quadrato 3×3). Le soluzioni del sudoku di difficoltà media al centro dell'illustrazione e degli altri che appaiono nell'articolo si possono trovare sul sito www.lescienze.it.

Jen Christiansen; fonte: Jean-Paul Delahaye

GLI ANTENATI DEL SUDOKU

Un sudoku è un tipo speciale di quadrato latino. I quadrati latini, che furono chiamati così da Eulero, matematico svizzero del Settecento, sono matrici $n \times n$ in cui compaiono n simboli in modo che lo stesso simbolo non compare mai due volte nella stessa riga o nella stessa colonna. Qui ne mostriamo due esempi. Un sudoku standard completato è un quadrato latino 9×9 che soddisfa il vincolo aggiuntivo che ognuno dei suoi nove sottoschemi contiene le cifre da 1 a 9.

Casella

1	2	3	4
2	3	4	1
3	4	1	2
4	1	2	3

Piccolo quadrato latino (n=4)

Sottogriglia

5	8	6	4	2	1	3	7	9
3	2	7	9	6	5	4	8	1
9	1	4	3	7	8	6	2	5
1	6	3	5	8	4	7	9	2
2	4	5	1	9	7	8	6	3
8	7	9	6	3	2	5	1	4
7	5	8	2	1	3	9	4	6
6	3	1	7	4	9	2	5	8
4	9	2	8	5	6	1	3	7

Quadrato latino che è anche un sudoku risolto (n=9)

Eulero

Jen Christiansen; fonte: Jean-Paul Delahaye (schemi); Emanuel Handmann, 1753, originale presso il Kunstmuseum di Basilea (Eulero)

numeri». La rivista depositò il nome, e così in Giappone gli imitatori usarono il nome «Number Place». Quindi un altro paradosso del sudoku è che i giapponesi lo chiamano con il suo nome inglese, mentre gli anglofoni lo chiamano col nome giapponese.

Il sudoku deve la sua successiva affermazione a Wayne Gould, un magistrato in pensione che oggi vive in Nuova Zelanda, che si imbatté nel gioco mentre visitava il Giappone nel 1997 e scrisse un programma per generare automaticamente griglie di sudoku. Alla fine del 2004 il «Times» di Londra accettò la proposta di Gould di pubblicare il gioco, seguito nel gennaio 2005 dal «Daily Telegraph». Da allora decine di giornali di tutto il mondo hanno iniziato a pubblicare sudoku, in qualche caso mettendolo addirittura in prima pagina come attrattiva promozionale. E sono spuntati libri e riviste, tornei, siti web e blog, tutti dedicati esclusivamente a questo gioco.

Un sudoku per tutti

Non c'è voluto molto perché i matematici si mettessero a giocare con il sudoku. Per esempio si sono subito chiesti quante griglie complete diverse, o «soluzioni», possono essere costruite. Chiaramente la risposta deve essere un numero inferiore a quello dei possibili quadrati latini, a causa del vincolo aggiuntivo delle sottogriglie.

Ci sono solo 12 quadrati latini di ordine 3 e 576 di ordine 4, ma ce ne sono 5.524.751.496.156.892.842.531.225.600 di ordine 9. La teoria dei gruppi, però, dice che uno schema può essere ricavato da un altro se è equivalente all'originale. Se sostituiamo sistematicamente ogni numero con un altro (per esempio, ogni 1 diventa 2, ogni 2 diventa 7 e così via), o se cambiamo di posto due righe o due colonne, gli schemi che otteniamo sono essenzialmente gli stessi. Se si contano solo le forme ridotte, il numero di quadrati latini di ordine 9 è 377.597.570.964.258.816 (risultato pubblicato su «Discrete Mathematics» nel 1975 da Stanley E. Bammel e Jerome Rothstein, all'epoca alla Ohio State University).

È stato piuttosto difficile stabilire quante griglie di sudoku possono esistere. Oggi solo l'uso della logica (per semplificare il problema) e del computer (per esaminare sistematicamente le tante possibilità) rende possibile una stima del numero di soluzioni, cioè griglie, diverse e valide: 6.670.903.752.021.072.936.960. Questo numero, ricavato da Bertram Felgenhauer del Politecnico di Dresda e da Franz Jarvis dell'Università di Sheffield, comprende tutti quelli derivati da una data griglia con operazioni elementari, ed è stato verificato diverse volte.

Se contiamo solo una volta le griglie che possono essere ridotte a configurazioni equivalenti, il numero si riduce a 5.472.730.538, poco meno del numero di esseri umani sulla Terra: gli appassionati del sudoku, dunque, non devono temere una penuria di giochi.

A una griglia di soluzione finale si può arrivare in più di un modo da una qualsiasi griglia iniziale (cioè da tutte le griglie incomplete la cui soluzione è una specifica griglia completa). Nessuno è ancora riuscito a determinare quanti diversi schemi iniziali esistano. Inoltre, uno schema iniziale del sudoku è interessante per un matematico solo se è minimo, cioè se rimuovere un singolo numero fa sì che la soluzione non sia più unica. Nessuno ha calcolato il numero di possibili schemi minimi esistenti, numero che darebbe la conta definitiva dei diversi sudoku, ma è una sfida che sicuramente sarà raccolta nel prossimo futuro.

C'è un altro problema di minimalità tuttora irrisolto: qual è il più piccolo numero di cifre che un creatore di sudoku deve inserire nello schema iniziale per garantire una soluzione unica? Sembra che la risposta sia 17. Gordon Royle della University of Western Australia ha raccolto più di 38.000 esempi con questa proprietà che non possono essere ricondotti l'uno all'altro con operazioni come scambi di righe o di colonne.

Gary McGuire della National University of Ireland a Maynooth, sta conducendo una ricerca per uno schema da 16 numeri, ma ancora non ne ha trovati e qualcuno inizia a pensare che non ne esistano. D'altro canto, Royle e altri in modo indipendente sono riusciti a trovare una griglia con 16 numeri che ha solo due soluzioni, e per ora nessun ricercatore ne ha trovate altre.

Qualcuno è vicino a dimostrare che nessuno schema valido di sudoku può avere solo 16 numeri? La risposta di McGuire è negativa. Se anche potessimo analizzare uno schema al secondo, osserva, ci vorrebbero 173 anni per studiarli tutti... Sfortunatamente non si può ancora fare, neppure con un computer velocissimo». Presto, secondo McGuire, su un computer abbastanza potente sarà possibile esaminare una griglia al minuto, ma a questo ritmo l'impresa richiederebbe 10.380 anni. «Persino distribuendo l'elaborazione su 10.000 computer ci vorrebbe circa un anno», aggiunge. E specifica: «Ci serve veramente un'innovazione teorica radicale per poter eseguire un'analisi di tutti gli schemi. O dobbiamo ridurre lo spazio della ricerca o dobbiamo trovare un algoritmo molto più efficiente».

I matematici conoscono la soluzione al problema opposto del minimo numero di cifre: qual è il massimo numero di cifre che non garantisce una soluzione unica? La risposta è 77. È molto facile vedere che con 80, 79 o 78 cifre se c'è una soluzione è unica, ma ciò non è detto nel caso di 77 cifre (*si veda il box in basso a p. 66*).

Soluzioni con il computer

Al di là delle domande su «quante sono», matematici e informatici si divertono a chiedersi che cosa i computer siano o non siano in grado di fare per la soluzione e la generazione di sudoku. Nel caso dei sudoku standard (9 × 9), è relativamente facile scrivere programmi in grado di risolvere tutte le griglie di partenza valide.

I programmi per trovare soluzioni possono usare vari metodi; il più comune è il *backtracking*, un approccio sistematico per tentativi ed errori in cui si provano soluzioni parziali e si modificano leggermente appena si riscontra che sono sbagliate.

L'algoritmo base del backtracking funziona così: il programma mette la cifra 1 nella prima casella vuota. Se la scelta è compatibile con le cifre già esistenti, passa alla seconda casella vuota, dove mette un 1. Quando trova un conflitto (il che può accadere molto presto), cancella l'1 appena inserito e inserisce al suo posto un 2 o, se non va bene, un 3 o la successiva cifra accettabile. Dopo aver inserito la prima cifra che non crea conflitto passa alla casella successiva e ricomincia da 1.

Se la cifra da modificare è 9 (che non può essere ulteriormente incrementata in un sudoku standard), il programma fa un passo indietro (*backtrack*) e incrementa di uno la cifra nella casella precedente (la penultima cifra inserita). Poi ricomincia ad avanzare finché non incontra un nuovo conflitto. (Qualche volta il programma fa diversi passi indietro prima di riprendere ad andare avanti.) In un programma ben scritto, questo metodo esplora tutte le pos-

FINO A DOVE SI PUÒ ARRIVARE?

Sembra che il minimo numero di cifre che devono comparire inizialmente in un sudoku 9×9 affinché abbia un'unica soluzione sia 17; qui sotto ne trovate un esempio. Un certo schema completo, noto ai conoscitori del sudoku come *Strangely Familiar* («stranamente familiare»), o SF, nasconde 29 schemi iniziali non equivalenti con 17 cifre: un numero insolitamente alto. Si pensava che SF fosse il miglior candidato ad avere uno schema iniziale con 16 cifre e un'unica soluzione, ma uno studio approfondito ha infranto questa speranza. In basso è mostrato l'unico sudoku noto che ha 16 cifre e solo due soluzioni; le griglie finali scambiano le posizioni degli 8 e dei 9.

Un sudoku con 17 cifre

	1							9
			3			8		
						6		
				1	2	4		
7		3						
5								
8			6					
				4			2	
			7				5	

Lo schema «Stranamente familiare»

6	3	9	2	4	1	7	8	5
2	8	4	7	6	5	1	9	3
5	1	7	9	8	3	6	2	4
1	2	3	8	5	7	9	4	6
7	9	6	4	3	2	8	5	1
4	5	8	6	1	9	2	3	7
3	4	2	1	7	8	5	6	9
8	6	1	5	9	4	3	7	2
9	7	5	3	2	6	4	1	8

Un sudoku con 16 cifre...

5		2				4		
			7	1				3
					4	6		
	7		2					
	1							
6					2			
				3			1	
4								

...con due soluzioni

5	6	2	3	8/9	9/8	4	7	1
8/9	4	9/8	7	1	6	2	5	3
1	3	7	4	2	5	8/9	9/8	6
3	5	8/9	1	9/8	4	6	2	7
9/8	7	4	2	6	3	1	8/9	5
2	1	6	8/9	5	7	3	4	9/8
6	9/8	1	5	4	2	7	3	8/9
7	2	5	6	3	8/9	9/8	1	4
4	8/9	3	9/8	7	1	5	6	2

Jen Christiansen; fonte: Gordon Royle/University of Western Australia

Qual è il numero di cifre più piccolo che si deve inserire nello schema iniziale di un sudoku per avere un'unica soluzione?

In sintesi/*Sudoku scientifico*

- Il sudoku non è solo un divertente gioco di logica, ma solleva anche una gran quantità di problemi interessanti per i matematici.
- Tra questi problemi ci sono: quanti sudoku possono essere costruiti? Qual è il minimo numero di cifre che devono essere presenti all'inizio perché si abbia un'unica soluzione? Il sudoku appartiene a una classe di problemi difficili noti come problemi NP-completi, che pongono ostacoli insuperabili al software.
- Gli esperti di enigmi hanno messo a punto numerosi metodi per affrontare i sudoku, e parecchie divertenti varianti del gioco.

METODI PER LA SOLUZIONE

Ecco alcuni metodi per provare a risolvere un sudoku. I metodi *1* e *2* sono i più semplici, e di solito vengono usati insieme (un po' l'uno e un po' l'altro). Purtroppo spesso non portano lontano, e quindi il giocatore può aiutarsi con il metodo *3* e, se non basta, con il metodo *4*, che funziona sempre ma non necessariamente in modo semplice. È anche possibile inventare metodi propri e provare i numerosi approcci descritti sul Web.

a

5		1					9	6
				9			5	
					5	2		7
4	9		1				7	
					7			
1	3						2	
3		4		5	9			
	2	8		7	1		4	
7	6	5	8	2				

b

Caselle «uniche» (*in blu*)

Caselle «forzate» (*in arancione*)

5		1					9	6
				9			5	
				1	5	2		7
4	9		1				7	
	5				7			
1	3	7					2	
3	1	4	6	5	9	7		2
9	2	8		7	1		4	
7	6	5	8	2				

c

5	478	1	2347	348	2348	348	9	6
268	478	2367	23467	9	23468	1348	5	1348
689	48	369	346	13468	5	2	138	7
4	9	26	1	368	2368	3568	7	358
268	58	26	234569	3468	7	1345689	1368	134589
1	3	67	4569	468	468	45689	2	4589
3	1	4	6	5	9	1678	168	128
9	2	8	36	7	1	3569	4	359
7	6	5	8	2	34	139	13	139

d

Jen Christiansen; fonte: Jean-Paul Delahaye

METODO 1

CASELLA «FORZATA»

Questo metodo fissa l'attenzione su una specifica casella. Eliminando dalle possibili soluzioni tutti i numeri che si trovano nella stessa colonna, riga o sottogriglia, si può vedere se rimane un'unica soluzione. Da un'analisi di questo tipo si vede che le caselle che contengono numeri arancioni nello schema *b* sono «forzate».

METODO 2

CASELLA «UNICA»

In questo metodo ci si concentra su un certo numero, per esempio il 5. Nella prima e nella terza colonna dello schema *a* il 5 è già presente, ma finora non ce n'è nessuno nella seconda colonna. Dove può essere il 5 di quella colonna? Non nelle prime tre caselle della seconda colonna, perché si trovano in una sottogriglia in cui è già presente. Non nella settima casella della colonna, perché anche la sua sottogriglia ha un 5. Quindi il 5 della seconda colonna si trova nella quarta, quinta o sesta casella della colonna. La quinta è l'unica libera, e quindi il numero va lì. Le caselle con numeri blu nello schema *b* sono caselle «uniche».

METODO 3

SEMPLIFICARE L'INSIEME DELLE SOLUZIONI POSSIBILI

Questa tecnica è molto potente, ma richiede matita e gomma. In ogni casella si scrivono tutte le possibili soluzioni con numeri molto piccoli o usando puntini le cui posizioni rappresentano i numeri da 1 a 9. Poi si applica la logica per eliminare alcune opzioni. Per esempio, lo schema *c* raffigura l'aspetto che avrebbe lo schema *a* se tutte le sue caselle fossero riempite meccanicamente, senza applicare i metodi *1* e *2*. Nella terza colonna gli insiemi delle possibili soluzioni (*primo ingrandimento*) per la seconda, terza, quarta, quinta e sesta casella sono, rispettivamente, {2,3,6,7}, {3,6,9}, {2,6}, {2,6} e {6,7}. La colonna deve contenere un 2 e un 6, e quindi questi due numeri devono trovarsi nelle due caselle le cui uniche soluzioni sono 2 e 6 (*cerchiati*). Di conseguenza, in questa colonna il 2 e il 6 non possono essere da nessun'altra parte e possono essere cancellati dalle altre caselle della colonna (*in rosso*). Lo spettro delle possibilità per la colonna si è ridotto a {3,7}, {3,9}, {2,6}, {2,6} e {7}. Ma non è tutto. Individuare la posizione del 7 ci dà a sua volta le posizioni del 3 e del 9 (*secondo ingrandimento*). Le possibilità finali sono {3}, {9}, {2,6}, {2,6} e {7}. Rimane una sola incertezza: le posizioni del 2 e del 6. La regola generale per semplificare le soluzioni possibili è la seguente: se in un insieme di soluzioni (per una riga, una colonna o una sottogriglia) si trovano *m* caselle che contengono complessivamente un sottoinsieme fatto di *m* numeri (non necessariamente tutti in ogni casella), le cifre di questo sottoinsieme possono essere eliminate dalle possibilità per le altre caselle dell'insieme più grande. Per esempio, in *d*, la situazione {2,3}, {1,3}, {1,2}, {1,2,4,5}, {3,5,7} può essere semplificata a {2,3}, {1,3}, {1,2}, {4,5}, {5,7} perché le caselle {2,3}, {1,3}, {1,2} provengono tutte dal sottoinsieme {1,2,3} e non hanno altri numeri.

METODO 4

TENTATIVI ED ERRORI

Applicando i metodi da *1* a *3* si possono risolvere molti sudoku. Ma quelli di livello «diabolico» spesso richiedono una procedura che va per tentativi ed errori. Quando resta un'incertezza si compie una scelta casuale e si applicano tutte le strategie come se fosse la decisione corretta. Se ci si imbatte in una soluzione impossibile (due numeri identici nella stessa colonna) si sa di aver effettuato una scelta scorretta. Per esempio, si potrebbe provare il 2 nella quarta casella della terza colonna dello schema *c*. Se si rivela sbagliata, si ricomincia dallo stesso punto, ma questa volta ponendo in quella casella il 6.

Purtroppo spesso è necessario provare diverse volte la strategia per tentativi ed errori, e bisogna essere pronti a fare dei passi indietro se si sono compiute scelte sbagliate. L'idea dietro a questo metodo è la stessa usata negli algoritmi di backtracking, che possono essere facilmente implementati nei programmi per computer ma che sono insopportabilmente faticosi per la mente umana. È singolare che il metodo più efficiente per una macchina sia il meno efficiente per un essere umano.

sibili ipotesi fino a trovare la soluzione, se esiste. E se esiste più di una soluzione, il che può succedere per un sudoku imperfetto, il programma le trova tutte.

Naturalmente sono possibili miglioramenti che accelerano la scoperta dell'unica soluzione. Uno molto apprezzato è detto *constraint propagation* (propagazione del vincolo): ogni volta che si inserisce un nuovo numero, il programma genera una tabella dei numeri ammissibili rimanenti per ogni casella vuota e considera solo i numeri di questa tabella.

Le tecniche di backtracking possono essere incorporate in programmi piuttosto piccoli. In particolare, per il sudoku sono stati scritti programmi molto brevi in Prolog, un linguaggio di programmazione che include un algoritmo di backtracking inventato da Alain Colmerauer e Philippe Roussel dell'Università di Marsiglia alla fine degli anni settanta.

Per i giocatori in carne e ossa, le tecniche di soluzione basate sul backtracking non sono praticabili, perché richiederebbero una pazienza sovrumana. Quindi gli esseri umani usano regole più varie e più furbe, e in genere ricorrono al metodo per tentativi ed errori solo come ultima possibilità. Alcuni programmi cercano di imitare almeno in parte i nostri metodi: sono più lunghi degli altri, ma funzionano altrettanto bene. I programmi che simulano il ragionamento umano sono utili anche per valutare la difficoltà degli schemi, che sono classificati da «facile» (quelli che richiedono solo tattiche semplici) a quello che molti chiamano «diabolico» (a causa della necessità di applicare regole logiche più complesse).

Una delle strategie adottate dagli scienziati per trovare la soluzione di un sudoku consiste nel ricondurre il problema a un problema di colorazione di un grafo, in cui due caselle «vicine» (corrispondenti a due vertici congiunti da un arco) non possono avere lo stesso colore e la tavolozza disponibile è formata da nove colori. Il grafo contiene 81 vertici, di cui alcuni colorati fin dall'inizio. Il problema, noto come «problema della colorazione», in realtà è piuttosto complesso, perché ogni schema 9 × 9 ha centinaia di archi. Ogni casella fa parte di una riga che comprende altre otto caselle, di una colonna che comprende altre otto caselle e di una sottogriglia che ne comprende ancora altre otto (di cui quattro sono già state contate nella colonna o nella riga). Quindi ognuna delle 81 caselle è «adiacente» ad altre 20 (8 + 8 + 4) caselle; si hanno così complessivamente 1620 coppie di caselle adiacenti, il che significa che nel grafo ci sono 810 archi (1620 diviso 2).

L'obiettivo apparentemente impossibile consiste nel progettare un programma che risolva sudoku di qualsiasi dimensione

Il fatto che i sudoku possano essere tradotti in termini di problema della colorazione è significativo per gli scienziati, perché questa proprietà collega il gioco a una classe di problemi importanti. In particolare, Takayuki Yato e Takahiro Seta, entrambi dell'Università di Tokyo, hanno dimostrato di recente che il sudoku appartiene alla categoria dei «problemi NP-completi», problemi che probabilmente non è possibile risolvere in un arco di tempo realistico. Uno degli esempi più noti è il «problema della 3-colorabilità», in cui ci si chiede se è possibile assegnare a ogni nodo di un grafo un colore scelto fra tre, in modo che a due nodi congiunti da un arco non sia mai assegnato lo stesso colore.

Nel caso del sudoku, l'obiettivo apparentemente impossibile consiste nel progettare un programma efficiente che risolva sudoku di qualsiasi dimensione, cioè qualsiasi schema della forma $n^2 \times n^2$, e non solo la versione standard $3^2 \times 3^2$ (9 × 9). Nessun programma in grado di risolvere qualsiasi schema funzionerebbe efficientemente, perché il tempo necessario per trovare una soluzione aumenta enormemente al crescere di n.

Se avessimo un algoritmo in grado di risolvere i sudoku, potremmo usarlo per ottenere un algoritmo in grado di crearli. Infatti, anche se i primi sudoku erano costruiti «a mano», oggi quasi tutti sono generati da programmi basati sull'approccio che sto per descrivere o su uno simile. Si dispongono a caso alcuni numeri in uno schema vuoto e si applica un algoritmo per la ricerca di soluzioni (per esempio il backtracking). Se lo schema ha un'unica soluzione, il programma si arresta. Se lo schema completato parzialmente non ammette soluzioni, si toglie un numero e il programma ricomincia. Se lo schema ha

VARIAZIONI SUL TEMA

Siete alla ricerca di qualcosa che vada al di là degli usuali sudoku diabolici? Nei giochi illustrati qui accanto si applicano le regole abituali del sudoku, ma con qualche trovata in più. In *a* le lettere delle parole MARTIN GARDNER, indimenticato autore di una rubrica di giochi matematici su queste pagine, sostituiscono i numeri, e le forme geometriche sostituiscono i sottoschemi quadrati. L'inventore ha chiamato il gioco «Du-Sum-Oh». In *b*, una stella con solo sei sottogriglie, alcune righe e colonne sono interrotte al centro, e le punte completano le righe più corte. In *c* i numeri di tre cifre delle prime due sottogriglie nelle righe contrassegnate dai segni aritmetici hanno come somma i numeri nella terza sottogriglia. In *d* i segni di maggiore e minore servono a trovare le posizioni delle cifre. In *e* bisogna inserire negli spazi vuoti i pezzi del domino che si trovano in basso. In *f* si sovrappongono tre schemi. Nel sito www.lescienze.it potete trovare le soluzioni a tutti questi schemi.

a

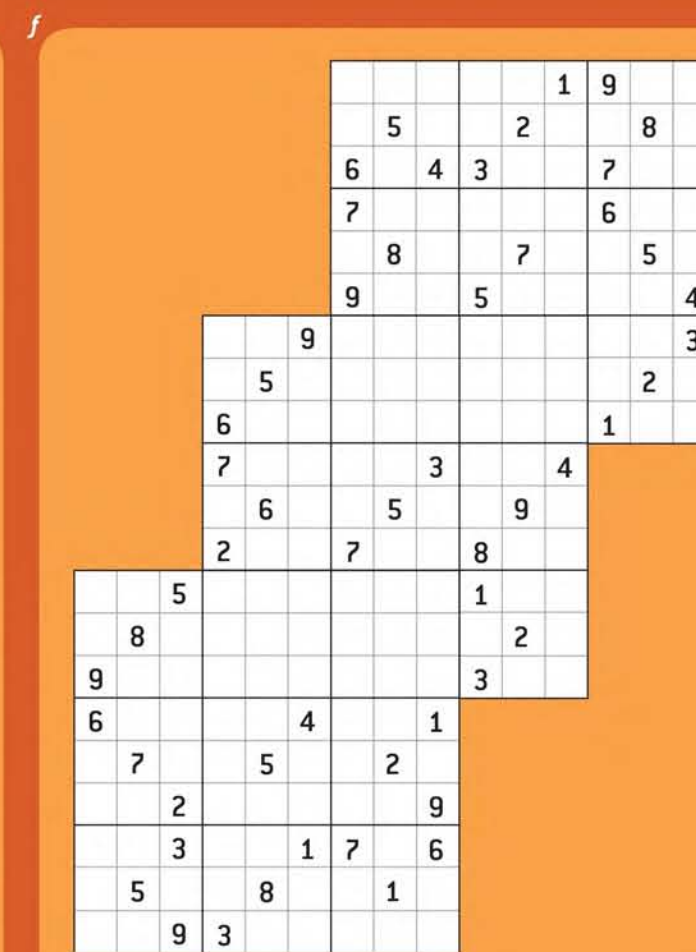
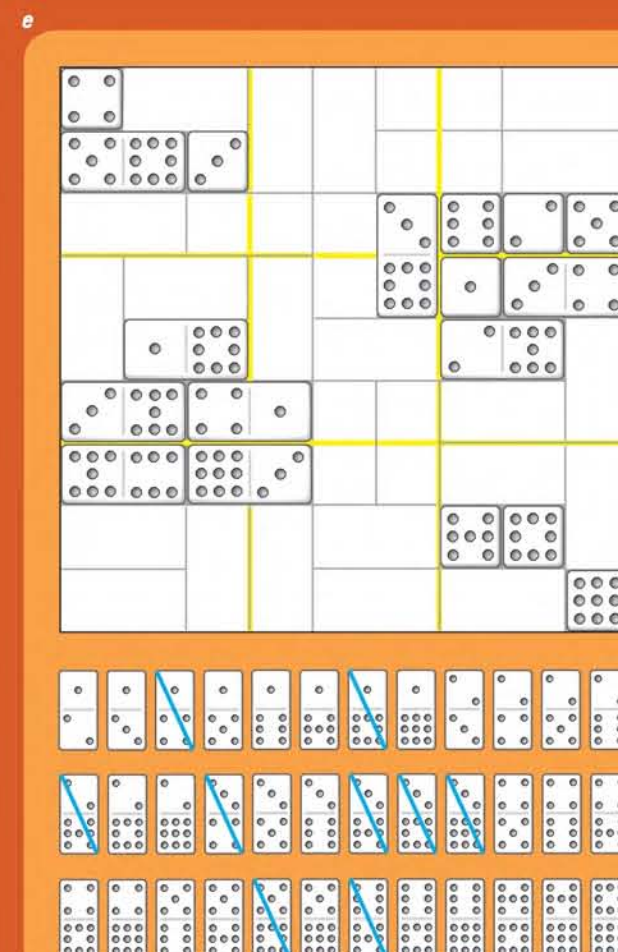
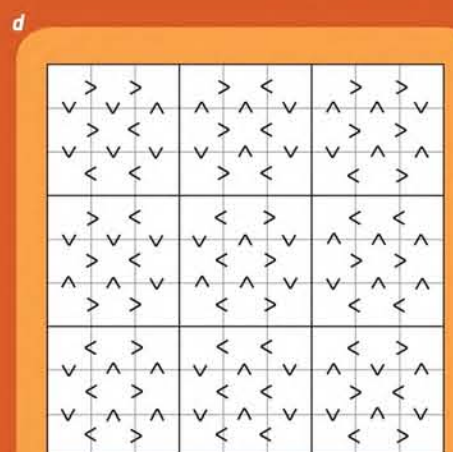
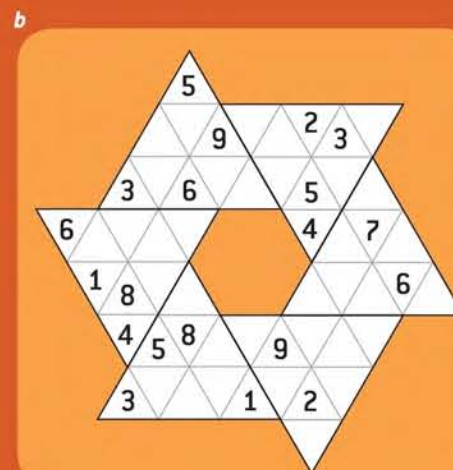
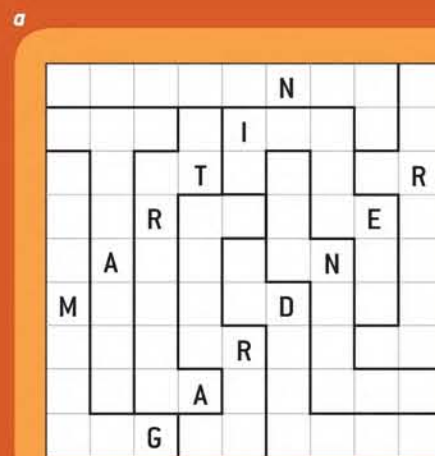
b

c

d

e

f

Jen Christiansen; fonte: Bob Harris, *In the Grand Time Sudoku and the Law of Leftovers*, A K Peters Ltd (in stampa) (*a*); © Nonzero/Olga Leontieva (*b*); Ed Pegg, Jr. (*c*); © Nonzero/Cihan Altay (*d*) Ed Pegg, Jr. (*e* e *f*)

più di una soluzione, se ne sceglie una e l'algoritmo aggiunge allo schema iniziale tutti i numeri che servono per garantire che la soluzione scelta sia unica.

Strategie umane

Gli appassionati che si divertono a risolvere il sudoku a mano possono scegliere tra molte tattiche, ma ci sono due approcci fondamentali che offrono un buon punto di partenza. Il primo consiste nel cercare tra le caselle vuote quelle con più vincoli: cioè quelle che appartengono a una riga, colonna o sottogriglia che sia già in buona parte riempita. Qualche volta, eliminando le cifre non valide (i numeri che già occupano caselle della stessa riga, colonna o sottogriglia), si scopre l'unico numero che va bene in quella casella; in ogni caso, questo metodo riduce di molto le opzioni.

Nel secondo approccio, si cerca dove può stare una data cifra in una certa colonna, riga o sottogriglia (per esempio, cercare gli unici posti in cui il 3 possa trovarsi nella quarta riga). Qualche volta c'è un'unica risposta, altre volte ci si accorge che è utile anche sapere semplicemente che il 3 può andare solo in due o tre posti specifici. Nel box alle pagine 64-65 ci sono ulteriori dettagli; visitando i siti web elencati qui sopra, si troverà una gran quantità di strategie, alcune delle quali hanno nomi fantasiosi come *swordfish* (pesce spada) e *golden chain* (catena d'oro).

Molti programmi che si possono trovare facilmente in rete generano schemi di difficoltà predefinita e aiutano a trovare la soluzione (senza risolvere il gioco al vostro posto!). Alcuni per esempio permettono di inserire nelle caselle dei segni provvisori e di cancellarli, rendendo così inutili gomma e matita, altri permettono persino di creare collegamenti tra caselle. Non sottovalutate questi programmi. Liberandovi da attività noiose come le cancellature, vi spingeranno verso maggiori sottigliezze e virtuosismi in questo gioco logico.

Una volta che vi sarete annoiati del sudoku tradizionale, potrete dare un'occhiata alle sue innumerevoli varianti: alcune sovrappongono diversi schemi, altre usano strutture diverse al posto delle sottogriglie quadrate e altre ancora usano vincoli aggiuntivi. L'attrattiva di queste varianti consiste nel costringere a esplorare nuove strategie. E gli appassionati che impiegano cinque minuti per risolvere un gioco tradizionale possono immergersi nelle delizie che si scoprono combinando caselle e numeri in versioni gigantesche del sudoku. Ma ora basta. Al prossimo sudoku!

IL MASSIMO PER PIÙ SOLUZIONI

Non è detto che 77 numeri siano sufficienti a garantire un'unica soluzione. Nonostante abbia solo quattro caselle vuote, questo schema ha due soluzioni: nelle prime due colonne i numeri 1 e 2 mancanti sono intercambiabili.

1 2	2 1	3
4	5	6
7	8	9
2 1	1 2	4

		3	4	5	6	7	8	9
4	5	6	7	8	9	1	2	3
7	8	9	1	2	3	4	5	6
		4	3	9	8	5	6	7
8	6	5	2	7	1	3	9	4
9	3	7	6	4	5	8	1	2
3	4	1	8	6	2	9	7	5
5	7	2	9	1	4	6	3	8
6	9	8	5	3	7	2	4	1

L'AUTORE

JEAN-PAUL DELAHAYE è professore di informatica all'Université des Sciences et Technologies di Lille, in Francia, e ricercatore presso il Laboratoire d'Informatique Fondamentale de Lille del Centre National de la Recherche Scientifique che ha sede presso la stessa università. Le sue ricerche si concentrano sulla teoria dei giochi computazionale (per esempio sul dilemma del prigioniero iterato), sulla teoria della complessità (complessità di Kolmogorov) e sulle applicazioni di queste teorie all'analisi genetica e, di recente, all'economia.

PER APPROFONDIRE

Il sito del primo campionato mondiale di sudoku, svoltosi a Lucca il 10-11 marzo 2006: http://www.wsc2006.com/eng/index.php.

Giochi matematici di Ed Pegg, Jr.: http://www.maa.org/news/mathgames.html.

La matematica del sudoku, di Sourendu Gupta: http://theory.tifr.res.in/~sgupta/sudoku/.

Matematica del sudoku, di Tom Davis: http://www.geometer.org/mathcircles.

Tecniche per il sudoku (SadMan Software): http://www.sadmansoftware.com/sudoku/techniques.htm.

Sudoku in generale: http://www.sudoku.com/howtosolve.htm.

Voce «sudoku» di Wikipedia: http://it.wikipedia.org/wiki/Sudoku.

Varianti del sudoku: http://www.sudoku.com/forums/viewtopic.php?t=995.

I COMPONENTI BIOLOGICI sono gli elementi di partenza di un approccio alle biotecnologie modellato sull'esempio dell'ingegneria elettronica.

Slim Films

L'INGEGNERIA DELLA VITA

come costruire una fabbrica biologica

a cura del Bio Fab Group
David Baker, George Church, Jim Collins, Drew Endy, Joseph Jacobson, Jay Keasling, Paul Modrich, Christina Smolke e Ron Weiss

Dai successi dell'ingegneria si possono apprendere principi e pratiche per trasformare le biotecnologie da una professione specializzata in un'industria matura

Anche se l'espressione «ingegneria genetica» si usa almeno da trent'anni e le tecniche del DNA ricombinante sono ormai un cardine della ricerca moderna, in realtà l'ingegneria ha ben poco in comune con il lavoro che i biotecnologi fanno di norma sulla materia vivente. Uno dei motivi è che gli strumenti disponibili per le costruzioni costituite da «componenti» biologici devono ancora raggiungere un livello di standardizzazione e di funzionalità paragonabile a quello di altri settori dell'ingegneria. Un altro motivo ha a che fare con i metodi e la mentalità propri della biologia, che però possono anch'essi essere profondamente influenzati dalla tecnologia. L'ingegneria elettronica, per esempio, ha subito una profonda trasformazione a partire dal 1957, quando Jean Hoerni e Robert N. Noyce della Fairchild Semiconductors, una piccola azienda con sede in quella che in seguito sarebbe diventata la Silicon Valley, inventarono la cosiddetta «tecnologia planare». Si trattava di un sistema grazie a cui i metalli e le sostanze chimiche potevano essere disposti in strati sottili e incisi all'interno di un wafer di silicio usando calchi chiamati fotomaschere. Questo nuovo approccio consentì agli ingegneri di produrre circuiti integrati in modo costante e pulito e di creare una grande varietà di tipologie di circuiti semplicemente modificando lo schema della maschera. Presto gli ingegneri poterono attingere a circuiti elementari già creati da altri per combinarli tra loro a formare circuiti complessi con uno spettro di applicazioni sempre più vasto.

Prima di allora costruire un circuito elettronico significava dover collegare uno per uno i transistor da cui era composto. Era un processo artigianale, dai risultati incerti, un vero e proprio collo di bottiglia per la neonata industria elettronica. La tecnologia planare, invece, ha continuato a svilupparsi permettendo ai progressi tecnologici di procedere alla sorprendente velocità descritta dalla famosa legge di Moore.

Questa combinazione di fattori tecnologici e metodologici nel progettare e fabbricare i circuiti integrati a semiconduttori – la cosiddetta *chip fab* – costituisce per l'ingegneria uno dei paradigmi di maggior successo di tutti i tempi, ed è un valido modello anche per un altro settore tecnologico nascente: la produzione di sistemi biologici.

Gli ingegneri genetici di oggi sono ancora nella fase in cui collegano a mano i componenti dei circuiti. Come ha osservato Tom Knight del laboratorio di intelligenza artificiale del Massachusetts Institute of Technology: «La mancanza di standard nelle tecniche di assemblaggio del DNA fa sì che ciascuna reazione di assemblaggio sia allo stesso tempo uno strumento sperimentale per la ricerca in corso, ma anche un esperimento di per sé».

Uniformando i metodi e i componenti dell'ingegneria biologica si potranno creare archivi di elementi progettati *ad hoc* compatibili tra loro e si potranno fabbricare i propri prodotti delegando all'esterno la produzione dei componenti. Se il momento della progettazione fosse separato da quello della produzione, gli ingegneri biologici sarebbero liberi di immaginare dispositivi sempre più complessi, e per gestire la complessità potrebbero usare i potenti strumenti tipici dell'ingegneria, per esempio il computer design. Per questo alcuni membri del nostro gruppo hanno iniziato a individuare e sviluppare le strumentazioni e le tecniche che potrebbero diventare le basi di una «biofabbrica», e a promuovere la nascita di una comunità di scienziati che applichino alla biotecnologia i migliori principi e le migliori pratiche dell'ingegneria.

La bioingegneria è un modo di pensare alle macchine biologiche esistenti e a come costruirne di nuove

In sintesi/*La fabbrica della vita*

- Il metodo di fabbricazione dei circuiti a semiconduttori è frutto dell'incontro tra una tecnica flessibile e affidabile, metodi standardizzati e librerie di componenti a cui i progettisti possono attingere.
- Grazie a questo sistema gli ingegneri hanno potuto creare dispositivi elettronici straordinariamente complessi e potenti, con vaste applicazioni.
- Un approccio simile darebbe agli ingegneri biologici la possibilità di concepire e costruire sofisticati dispositivi costituiti da componenti biologici.
- Alcune tecnologie e tecniche per la biofabbricazione sono state sviluppate e sono già in uso. Ora si cerca di affrontare la questione della sicurezza e di portare i biologi ad assumere una mentalità da ingegneri.

Elementi di qualità

Se i transistor sono i componenti di base dei circuiti elettronici, i loro equivalenti biologici sono i geni: lunghi filamenti di DNA accuratamente ordinati. Per costruire circuiti genetici per dispositivi biologici avanzati, quindi, occorre sintetizzare lunghe sequenze di DNA in modo veloce e affidabile, e a un prezzo ragionevole.

Vent'anni fa Marvin H. Caruthers dell'Università del Colorado a Boulder, basandosi su studi precedenti, elaborò un sistema per sintetizzare singoli filamenti di DNA sfruttandone le proprietà chimiche naturali. Il DNA è composto da nucleotidi che si distinguono l'uno dall'altro per il tipo di subunità, detta base, che contengono: adenina (A), citosina (C), guanina (G) e timina (T). In base alle loro affinità, le basi si accoppiano – A con T e C con G – a formare i pioli di quella molecola a forma di scala a chiocciola che è il DNA. I vari gruppi chimici creano i legami sia tra le basi accoppiate sia tra i nucleotidi adiacenti di ciascun filamento.

Il metodo di Caruthers, noto come metodo delle fosforamiditi in fase solida, è tuttora alla base di moltissime tecniche di sintesi commerciali. Si parte con un singolo nucleotide ancorato a un supporto solido, come una sferetta di polistirolo, sospeso in un liquido. Quando è esposta a un acido, la base del nucleotide ancorato diventa pronta per formare un legame chimico con un altro nucleotide presente in soluzione. Anche questo secondo nucleotide viene esposto a un acido, legandone così un terzo e allungando la catena nascente. La ripetizione di questo ciclo consente di sintetizzare qualunque tipo di sequenza con una frequenza di errore di circa una base su 100.

Spesso però gli ingegneri genetici hanno l'esigenza di produrre strutture molto più lunghe di 100 basi. Un complesso genico può contare diverse migliaia di basi, mentre il genoma di un organismo anche molto piccolo, come un batterio, arriva a diversi milioni di basi. Molti ricercatori del nostro settore, desiderosi di metodi di sintesi con rese più alte e frequenze di errore più basse, si sono quindi rivolti alla natura in cerca di aiuto.

Negli organismi viventi il macchinario biologico composto da enzimi come la polimerasi è in grado di costruire e riparare molecole di DNA con una velocità che arriva a 500 basi al secondo e una frequenza di errore di una base su un miliardo. Il che significa una produttività (cioè la resa divisa per la frequenza di errore) tre miliardi di volte superiore rispetto ai migliori strumenti esistenti per la sintesi del DNA, che inseriscono una base ogni 300 secondi. Inoltre la copia di un tratto di DNA molto lungo, come il genoma di un batterio, è eseguita da diverse molecole di polimerasi che lavorano in parallelo, riuscendo a produrre circa cinque milioni di basi in 20 minuti.

George Church, uno degli autori di questo articolo, ha pensato di emulare il lavoro in parallelo delle polimerasi, introducendo alcune modifiche alla tecnologia dei *microarray*. Questi ultimi sono supporti rigidi a cui sono ancorati i cosiddetti oligonucleotidi, molecole di DNA a filamento singolo della lunghezza di 50-70 basi. Gli oligonucleotidi sono sintetizzati tutti contemporaneamente sulla superficie del microarray usando il metodo delle fosforamiditi, distribuiti secondo uno schema reticolare con una densità che arriva fino a un milione di punti per centimetro quadrato. Alla tecnologia tradizionale abbiamo aggiunto l'uso del *linker*, cioè di una breve sequenza di DNA posta alla base dell'oligonucleotide che può essere tagliata, così da consentire il distacco dell'oligonucleotide dal microarray. Ciascun punto nel nostro microarray sperimentale è grande più o meno 30 micrometri e contiene circa dieci milioni di molecole di DNA.

COMPONENTI VIVENTI

I batteri programmati si dispongono secondo uno schema ad anello in risposta a segnali prodotti da colonie di cellule «emittenti» che appaiono rosse (*a sinistra*), gialle (*al centro*) o rosa (*a destra*). «Prati» di cellule «riceventi» ingegnerizzate di *Escherichia coli* riconoscono le sostanze prodotte dalle cellule emittenti poste in mezzo a loro. I batteri riceventi producono proteine fluorescenti dai colori diversi a seconda della distanza, rilevata con la concentrazione della sostanza emessa.

Si possono creare schemi sempre più complessi modificando la posizione iniziale delle colonie emittenti. Questo esempio di sistema multicellulare complesso può essere usato per comprendere i meccanismi di segnalazione intercellulare, per esempio durante lo sviluppo di un organismo. La stessa tecnologia può essere applicata ai sensori, alla produzione artificiale di tessuti su tre dimensioni e alla biofabbricazione di materiali da parte di organismi programmati.

Cortesia Ron Weiss/Princeton University (*in alto*); Aaron Chevalier/Università del Texas ad Austin/Università della California a San Francisco (*sotto*)

A questi filamenti abbiamo dato il nome di oligonucleotidi da costruzione perché le loro sequenze parzialmente sovrapposte consentono di assemblarli e costruire tratti di DNA più lunghi, per esempio un gene intero. Gli oligonucleotidi che contengono errori devono però essere eliminati. A questo scopo, abbiamo elaborato due diversi sistemi di correzione degli errori.

Il primo usa lo stesso metodo di sintesi del microarray per produrre gli oligonucleotidi da selezione, con sequenze complementari agli oligonucleotidi da costruzione. Gli oligonucleotidi da selezione, una volta sintetizzati, sono rilasciati dal supporto solido e fatti scorrere sulla superficie del microarray a cui sono legati gli oligonucleotidi da costruzione. Qui si ibrideranno con gli oligonucleotidi complementari secondo le regole dell'appaiamento tra le basi, formando molecole di DNA a doppio filamento. A questo punto saremo in grado di identificare ed eliminare le molecole di DNA da costruzione che, avendo errori nella sequenza, non si sono ibridate o lo hanno fatto con errori grossolani. Si noti che anche gli oligonucleotidi da selezione possono contenere errori, visto che sono costruiti nello stesso modo degli oligonucleotidi da costruzione, ma la probabilità che una sequenza erronea in uno dei due gruppi trovi una complementarità nell'altro è molto bassa. Quindi, usare un gruppo di oligonucleotidi come «correttore di bozze» per l'altro gruppo è un approccio efficace, che ci consente di produrre oligonucleotidi con una media di errori di una base ogni 1300.

I sistemi biologici hanno tutto l'interesse a riprodurre se stessi in modo accurato, e in effetti il nostro secondo metodo di correzione è preso in prestito dalla natura. Modrich, un altro degli autori, dieci anni fa ha studiato a fondo i dettagli del processo di correzione e lo ha chiamato «sistema MutS, L, H». Quando l'ibridazione di due filamenti di DNA porta a un appaiamento imperfetto tra le basi A-T e G-C, nel punto del disappaiamento la molecola assume una forma diversa dalla classica doppia elica. MutS è una proteina naturale che riconosce e lega queste imperfezioni e, in alcuni casi, recluta MutL o MutH. Jacobson, un altro membro del Bio Fab Group, insieme con Peter Carr del Massachusetts Institute of Technology (MIT), ha usato questo sistema per raggiungere frequenze di un errore ogni 10.000 basi di DNA sintetico: un'accuratezza sufficiente a produrre piccoli complessi genici.

I SALUTI DAI BATTERI brillano su questo congegno che nel 2004 ha partecipato alla competizione dell'iGEM ed è stato prodotto all'Università del Texas ad Austin. Gli scienziati hanno incorporato in cellule di *Escherichia coli* diverse componenti genetiche che ricevono la luce e producono colore, in modo da trasformare un biofilm in una pellicola biologica che riproduce immagini impresse dalla luce sulla sua superficie. In continuità con la tradizione della programmazione informatica, il primo messaggio prodotto dalla macchina è stato: «Hello World».

Queste tecnologie – la sintesi in parallelo di oligonucleotidi rilasciabili e la correzione degli errori – ci consentono di assemblare lunghe molecole di DNA abbastanza esatte, in modo molto più rapido e meno costoso di quanto sia mai stato possibile finora. Si può quindi ritenere che queste procedure saranno le tecniche di base per la biofabbricazione, e che potranno portare a miglioramenti costanti nel tempo, proprio come la litografia per i circuiti a semiconduttori. A questo punto siamo liberi di concentrarci su quello che verrà costruito nella fabbrica.

La natura riveduta e corretta

Uno dei primi obiettivi che ci prefiggiamo è quello di usare una biofabbrica per esplorare nuove strategie di lotta alle malattie. Nei laboratori di Keasling e Baker, entrambi del nostro gruppo, si cercano nuove cure per la malaria e per l'infezione da HIV, due delle malattie che più affliggono l'umanità. Gli strumenti terapeutici che si cercano sono diversi, ma in entrambi i casi il lavoro si basa sulla sintesi di sequenze di DNA lunghe e accurate. Quindi i nostri progetti offrono un esempio di come l'approccio ingegneristico della biofabbrica possa cambiare il modo in cui gli scienziati biomedici procedono alla ricerca di nuove cure.

Nel caso della malaria esiste già un trattamento in grado di eliminare il parassita dal corpo di una persona infetta. Si tratta di un «C-15 sesquiterpene», una piccola molecola comunemente nota come artemisina che si estrae dall'*Artemisia annua*, un arbusto diffuso soprattutto nella Cina settentrionale. Le piante però la producono in quantità troppo basse perché possa essere impiegata su larga scala a costi accettabili. Per questo il gruppo di Keasling ha dedicato gli ultimi cinque anni a cercare di copiare la serie completa di geni che porta alla produzione di artemisina nella pianta, e a inserirla nel lievito per produrla in grandi quantità.

Una volta introdotta nel lievito, la serie dei geni può anche essere modificata per fare in modo che lavori con più efficienza rispetto alla pianta nativa. Per ora siamo riusciti a ridisegnare alcuni importanti sottoinsiemi di geni, noti come «via del mevalonato», per produrre un precursore dell'artemisina con un rendimento 100.000 volte superiore rispetto a quello che il processo naturale produce nei batteri. Per un ulteriore incremento dell'efficienza, che arrivi a rendere la cura disponibile su vasta scala, sarà necessario riprogettare l'intero percorso genico dell'artemisina con un approccio integrato.

Il sistema completo consiste di nove geni della lunghezza media di circa 1500 basi ciascuno. Qualunque nuova versione, quindi, conterrà approssimativamente 13.000 basi. Sarebbe inoltre utile poter costruire più di una variante per ogni singolo gene per provare diverse combinazioni e vedere quale funziona meglio. La sintesi di due sole varianti per ogni gene porta a 512 sequenze, per un totale di circa sei milioni di nucleotidi. Si tratterebbe di un obiettivo ambizioso se si usassero le tecniche ordinarie per la sintesi del DNA, ma ricorrendo alla tecnica del microarray è sufficiente una sintesi su un unico supporto.

La stessa tecnologia che rende possibile la sintesi di sistemi genici su vasta scala può anche essere impiegata per la creazione di nuove proteine, per esempio nuovi catalizzatori per reazioni di sintesi o per la riqualificazione ambientale oppure enzimi altamente specifici per la terapia genica o per la distruzione dei patogeni. Il gruppo di Baker sta sviluppando metodi computazionali per disegnare proprio le strutture di queste nuove proteine, due delle quali mimano i tratti essenziali presenti sulla superficie del virus HIV e sono già alla prova come potenziali vaccini.

Il problema è che i modelli informatici non sono abbastanza avanzati da garantire che le proteine progettate abbiano davvero le funzionalità desiderate. I computer possono però generare decine o centinaia di strutture candidate da mettere poi alla prova. Trasformare tutte queste strutture in sequenze geniche richiederebbe la sintesi di centinaia di migliaia di basi, un compito che per la tecnologia corrente è impegnativo e costoso, ma che è senz'altro alla portata della prima generazione di biofabbriche.

I progetti appena descritti per la sintesi di DNA e proteine da usare contro la malaria e l'HIV illustrano un approccio reso possibile dalla tecnologia della biofabbrica che può essere applicato anche a molte altre malattie, comprese le nuove minacce emergenti. Per esempio, mettendo insieme il sequenziamento del DNA a basso costo e ad alta velocità (si veda *Genoma per tutti*, di George M. Church, in «Le Scienze» n. 449, gennaio 2006) con le capacità di sintesi della biofabbrica, si potrebbero caratterizzare con eccezionale rapidità nuovi virus come la SARS o nuovi ceppi influenzali, e si potrebbero preparare vaccini basati sulle proteine virali molto più rapidamente di quanto sia possibile oggi.

Ma una biofabbrica è molto di più che una serie di tecnologie di sintesi più veloci: è un nuovo modo di pensare ai macchinari biologici esistenti e a costruirne di nuovi, che prende in prestito sia il linguaggio sia i metodi dell'ingegneria.

Una delle sfide del futuro sarà quella di portare sempre più biologi a pensare come gli ingegneri e viceversa

I mattoni biologici

Nel 2000 Michael Elowitz e Stanislas Leibler, allora alla Princeton University, come anche Collins (uno degli autori) e i suoi colleghi Tim Gardner e Charles Cantor dell'Università di Boston, hanno costruito i primi elementi di base – un oscillatore ad anello e un interruttore a leva – per un circuito costituito da componenti biologici. Gli scienziati sapevano da almeno 25 anni che gli organismi usano circuiti di questo tipo per regolare i propri geni, ma questi due gruppi sono stati primi a produrre con successo «circuiti biologici artificiali» funzionanti.

Il significato di questo termine è illustrato molto bene dall'oscillatore ad anello di Elowitz e Leibler, che inizialmente fu prodotto nel tentativo di costruire un orologio biologico sintetico che aiutasse a capire meglio gli orologi naturali usati dai sistemi biologici. Il circuito di base consisteva in un anello di DNA, chiamato plasmide, contenente tre geni: *tetR*, *lacI*, e λ *cI*, che codificano rispettivamente per le proteine TetR, LacI, e λ cI. Ogni volta che un gene è tradotto in proteina, l'enzima polimerasi si deve legare a una regione del DNA, chiamata «promotore», situata a monte del gene stesso; solo allora la polimerasi può trascrivere il gene in RNA messaggero, che è poi tradotto in proteina. Se la polimerasi non riesce a legarsi al promotore, il gene non viene tradotto e la proteina non viene prodotta.

Il sistema di Elowitz e Leibler è stato messo a punto in modo che le proteine per cui codificavano i tre geni del circuito legassero l'uno il promotore dell'altro, selettivamente. Quindi la proteina LacI lega il promotore di *tetR*, mentre la proteina λ cI lega il promotore di *lacI* e la proteina TetR lega il promotore del gene λ *cI*. Queste relazioni fanno in modo che la proteina prodotta da un gene blocchi il legame della polimerasi a un altro gene. La produzione delle tre proteine avviene quindi in un ciclo oscillatorio: l'abbondanza di proteina LacI reprime l'attività del gene *tetR*; la conseguente assenza della proteina TetR permette l'espressione del gene λ *cI*, il cui prodotto reprime la produzione di LacI.

Se uno dei prodotti proteici di questo ciclo è legato anche al funzionamento di un gene che produce una proteina fluorescente e l'intero circuito è inserito in un batterio, si può osservare l'oscillazione del meccanismo grazie al fatto che il batterio si accende e si spegne a intermittenza come la freccia di un'auto. L'ultima versione dell'interruttore genetico del gruppo di Collins, invece, può essere usata per programmare i batteri in modo che rilevino i danni del DNA cellulare e comunichino questa scoperta disponendosi a formare un «prato» verde fluorescente, cioè un biofilm.

La cosa più straordinaria di questi circuiti biologici sintetici è forse il fatto che sono identici, come funzione, ai circuiti costruiti

PIÙ SICUREZZA PER SINTESI

L'esplorazione di tutte le nuove opportunità che la biofabbrica potrebbe offrire alla medicina, alla produzione di nuovi materiali e di sensori, al recupero di suoli inquinati, alla produzione energetica, è solo all'inizio. Ma come tutte le imprese comporta dei rischi. Una caratteristica dei sistemi biologici è loro capacità di evolversi e di replicarsi, fenomeno che suscita comprensibili preoccupazioni sul fatto che i dispositivi biologici possano causare danni, intenzionali o meno.

Nel 1975 ad Asilomar, in California, si tenne una conferenza per discutere di preoccupazioni simili in merito alla tecnologia del DNA ricombinante che all'epoca era nuova. Per la prima volta gli scienziati avevano la possibilità di estrarre un singolo gene da un organismo e inserirlo in un altro, dando vita a combinazioni genetiche non necessariamente presenti in natura. Oggi quella possibilità è uno strumento essenziale in qualunque laboratorio di biologia molecolare, anche grazie alle decisioni prese alla fine della conferenza di Asilomar che attenuarono le paure sull'uso del DNA ricombinante.

In un certo senso, quindi, le questioni sollevate riguardo alle nuove tecnologie per la biofabbricazione non sono nuove, ma la nostra comunità continua a discuterle. Un gruppo di scienziati e di bioeticisti ha affrontato le implicazioni della genomica di sintesi durante la conferenza Synthetic Biology 2.0, tenutasi in maggio all'Università della California a Berkeley. Le conclusioni, come i risultati di uno studio di 15 mesi, finanziato dalla Alfred P. Foundation, su rischi, benefici e precauzioni necessarie sono disponibili on line sul sito www.syntheticbiology.org.

Per ora gli scienziati possono adoperare le precauzioni consuete di un laboratorio che lavora in sicurezza biologica e osservare gli stessi codici etici che hanno funzionato per trent'anni. Naturalmente è facile assicurarsi che ricercatori responsabili si comportino in modo responsabile. Ma la preoccupazione è che un giorno la capacità di sintetizzare il DNA abbia una diffusione così ampia da consentire ai criminali di creare, per esempio, nuovi patogeni mortali. Questo pensiero ha spinto Church a proporre un sistema di monitoraggio, che potrebbe prevedere la registrazione di tutti coloro che lavorano con la biologia sintetica – così come al momento sono registrati dal governo degli Stati Uniti tutti i ricercatori che hanno a che

Sam Ogden

LE FIALE DI MATTONI BIOLOGICI riprese nella foto contengono frammenti di DNA da impiegare per scopi di ingegneria biologica. Queste fiale sono sicure per l'uso anche nei laboratori che hanno il livello di biosicurezza più basso.

fare con i cosiddetti reagenti speciali – o una sorveglianza sull'acquisto degli organismi di sintesi, della strumentazione e dei materiali precursori.

Un'altra possibilità è che la biofabbrica possa rappresentare essa stessa un sistema di massima sicurezza grazie al raffinato controllo che consente di esercitare. La maggior parte delle applicazioni descritte non richiedono che gli organismi sintetici siano esposti all'ambiente, ma nel caso fosse necessario gli organismi potrebbero essere creati con una codifica genetica diversa da quella naturale, rendendo impossibile lo scambio genico con altre forme di vita. Un dispositivo biologico sintetico potrebbe essere progettato per autodistruggersi dopo un certo numero di divisioni cellulari, oppure essere costruito in modo che la sua vita dipenda da un composto chimico non presente nell'ambiente. In ogni mattone biologico si potrebbe imprimere una filigrana genetica per identificare e tracciare gli organismi fabbricati con esso. In altre discipline dell'ingegneria la capacità di costruire meccanismi di alta precisione permette di accrescere la sicurezza, per esempio nei sistemi di controllo di volo a ridondanza tripla presenti sugli aerei. Noi riteniamo che questo possa dimostrarsi vero anche per i sistemi biologici sintetici costruiti in una biofabbrica. (*Bio Fab Group*)

dagli ingegneri elettrici per testare i nuovi processi di costruzione dei circuiti integrati semiconduttori. Gli ingegneri sanno che i componenti di base come l'oscillatore o l'interruttore sono completi dal punto di vista della logica. Poter costruire questi semplici componenti in modo accurato e affidabile apre le porte alla possibilità di progettare e fabbricare circuiti ben più complessi. E quando gli ingegneri biologici potranno contare su questi elementi di base potranno affrontare progetti più complessi come sistemi multicellulari, strutture a due o tre dimensioni o dispositivi con una funzione non biologica.

Weiss, altro componente del gruppo, ha recentemente prodotto un prototipo di sistema multicellulare che potrebbe essere impiegato, per esempio, nel rilevare la presenza di esplosivi o altre sostanze, comunicando poi il risultato attraverso un segnale visibile (*si vedi il box a p. 71*). Questa macchina consente di programmare milioni di cellule batteriche dando loro istruzioni e protocolli sia per comunicare tra loro durante l'esecuzione dei compiti sia per emettere i segnali luminosi in molti modi diversi.

Ispirato da questi primi esempi Endy, insieme a Knight e a Randy Rettberg del MIT, sta curando un registro di componenti biologici, il Registry of Standard Biological Parts, simile a quello di cui dispongono i progettisti di microchip. Il registro potrebbe aiutare a realizzare un ampio spettro di progetti di costruzioni biologiche, e la nostra speranza è che anche altri vi contribuiscano. Finora il registro contiene più di 1000 «mattoni biologici», alcuni analoghi a componenti elettronici, per esempio elementi che ricevono un certo segnale di input o emettono un output. Abbiamo anche definito un segnale portante standard, polimerasi per secondo o PoPS, simile alla corrente elettrica che scorre in un filo metallico di collegamento tra due componenti elettronici, in modo che gli ingegneri della biofabbrica possano combinare più facilmente e riutilizzare i singoli dispositivi.

Per dimostrare la forza di questo approccio e coltivare questo nuovo campo, il gruppo del MIT ha organizzato nel 2003 il primo corso di «ingegneria per biofabbrica» con componenti biologici. Le lezioni si sono presto trasformate in una competizione annuale a cui quest'estate partecipano le squadre di più di 30 università. Nella sua pur breve esistenza l'ambiente dell'International Genetically Engineered Machine (iGEM) ha già prodotto una quantità di congegni cellulari sorprendenti, tra cui biofilm che registrano e riproducono una fotografia.

Un'altra realizzazione dell'iGEM messa a punto da Smolke, Collins e Church è un congegno in grado di contare su base digitale usando una serie di segmenti di DNA. Bastano 20 «bit» di DNA per contare e segnalare più di un milione (2^{20}) di diversi stati cellulari. Si può incorporare questa tecnologia in un sensore che a sua volta potrebbe essere connesso a percorsi metabolici modificati, come quello per l'artemisina nella versione ottimizzata dal gruppo di Keasling, che farebbe compiere un salto in avanti alla produzione del farmaco.

Costruire la biologia sintetica

Quando abbiamo iniziato a lavorare alla costruzione della biofabbrica non esisteva una tecnica consolidata per produrre lunghe strutture di DNA in modo accurato, rapido e poco costoso. Oggi questa tecnica esiste, ed è solo una tra le tante presenti nella cassetta degli attrezzi, sempre più ricca, dell'ingegneria biologica. Sono ormai all'orizzonte i primi meccanismi biologici elaborati e modellati al computer, che solo alla fine prendono corpo sotto forma di strutture biologiche, proprio come i circuiti al silicio sono prima progettati e poi incisi.

Come nella costruzione di circuiti elettronici, questo approccio ha il vantaggio aggiuntivo di ottimizzare le interazioni tra le diverse componenti e di prevedere la presenza di errori, capacità che diventerà sempre più utile via via che aumenterà la complessità dei sistemi costruiti. Un altro vantaggio del progettare in astratto è che un ingegnere biologico non ha bisogno di costruirsi ogni singola parte a partire da zero e non deve nemmeno conoscerne il funzionamento interno, gli basta sapere che è affidabile.

Gli studenti che partecipano a iGEM potrebbero rappresentare la prima generazione di biologi-ingegneri che fin dall'inizio della carriera sono stati abituati a pensare a se stessi come appartenenti a entrambe le categorie. Un'altra sfida per il futuro sarà quella di portare sempre più biologi a pensare come gli ingegneri (e convertire più ingegneri alla biologia), specialmente per quanto riguarda la condivisione dei componenti. Finora le biotecnologie sono state caratterizzate da gruppi di lavoro isolati, che si occupano di sviluppare applicazioni con un unico, preciso scopo, come la produzione di un farmaco. Il futuro della tecnologia in campo biologico, invece, renderà necessario il contributo di molti gruppi diversi a un sistema unico. La nostra speranza è che la costruzione di una fabbrica della biologia faciliti questa tendenza e stimoli progressi rivoluzionari come quelli raggiunti dall'industria dei semiconduttori. ■

GLI ESORDI DELLA BIOFABBRICA

Un piccolo numero di aziende e di organizzazioni applica già i principi e gli strumenti dell'ingegneria alla produzione biologica commerciale: la biofabbrica si avvicina alla realtà.

AZIENDA	OBIETTIVO
BioBricks Foundation, Cambridge, Massachusetts	Promozione di strumenti, standard e componenti per l'ingegneria biologica, ad accesso libero.
Blue Heron Biotecnology, Bothell, Stato di Washington	Sintesi del DNA.
Amyris Biotechnologies, Emeryville, California	Modifica delle vie metaboliche per la produzione di farmaci nei microrganismi.
Codon Devices, Cambridge, Massachusetts	Costruzione di dispositivi biologici.
Foundation for Applied Molecular Evolution, Gainesville, Florida	Generazione di nuove proteine e di nuovi materiali.
Synthetic Genomics, Rockville, Maryland	Ingegneria per la produzione di carburante da microrganismi.

GLI AUTORI

I membri del BIO FAB GROUP sono DAVID BAKER dell'Università di Washington, GEORGE CHURCH della Harvard Medical School, JIM COLLINS dell'Università di Boston, DREW ENDY e JOSEPH JACOBSON del Massachusetts Institute of Technology, JAY KEASLING dell'Università della California a Berkeley, PAUL MODRICH della Duke University, CHRISTINA SMOLKE del California Institute of Technology, e RON WEISS della Princeton University. Sono amici, colleghi e a volte collaboratori, come nel caso della scrittura di questo articolo, che hanno firmato come gruppo perché le diverse competenze di ciascuno, e quindi i loro contributi al progetto della biofabbrica, impersonano la natura interdisciplinare dell'ingegneria biologica. Tutti gli autori sono inoltre consulenti scientifici della Codon Device di Cambridge, in Massachusetts, la prima impresa commerciale che si è lanciata nell'applicazione dei principi di ingegneria alla biologia di sintesi. Church, Endy, Jacobson e Keasling sono tra i fondatori dell'azienda. Endy è fondatore anche dell'organizzazione no-profit BioBricks Foundation e Keasling ha creato la Amyris Biotechnologies.

PER APPROFONDIRE

GIBBS W.W., *Vita sintetica*, in «Le Scienze» n. 430, giugno 2004.

ENDY D., *Foundation for Engineering Biology*, in «Nature», Vol. 438, pp. 449-453, 24 novembre 2005.

CHURCH G., *Let Us Go Forth and Safely Multiply*, in «Nature», Vol. 438, p. 423, 24 novembre 2005.

ENDY D., DEESE I., THE MIT SINTHETYC BIOLOGIC WORKING GROUP e WADEY C., *Adventures in Synthetic Biology*. Un libro a fumetti che si può sfogliare su Internet all'indirizzo http://openwetware.org/wiki/Adventures.

Registry of Standard Biological Parts: http://parts.mit.edu.

Viaggiare gratis nel sistema solare

Sorprendenti trucchi matematici permettono di guidare un veicolo spaziale su grandi distanze con un consumo di combustibile ridotto o addirittura nullo

di Shane D. Ross

PER GIUNGERE A DESTINAZIONE, le sonde consumano una quantità di carburante che può arrivare al 50 per cento del loro peso. Ma la presenza di «autostrade planetarie» a basso consumo di energia potrebbe rivoluzionare il futuro dei viaggi spaziali. Nell'immagine, rappresentazione artistica del Mars Reconnaissance Orbiter, l'ultima sonda NASA inviata verso Marte, lanciata il 12 agosto 2005 ed entrata in orbita alla fine di marzo.

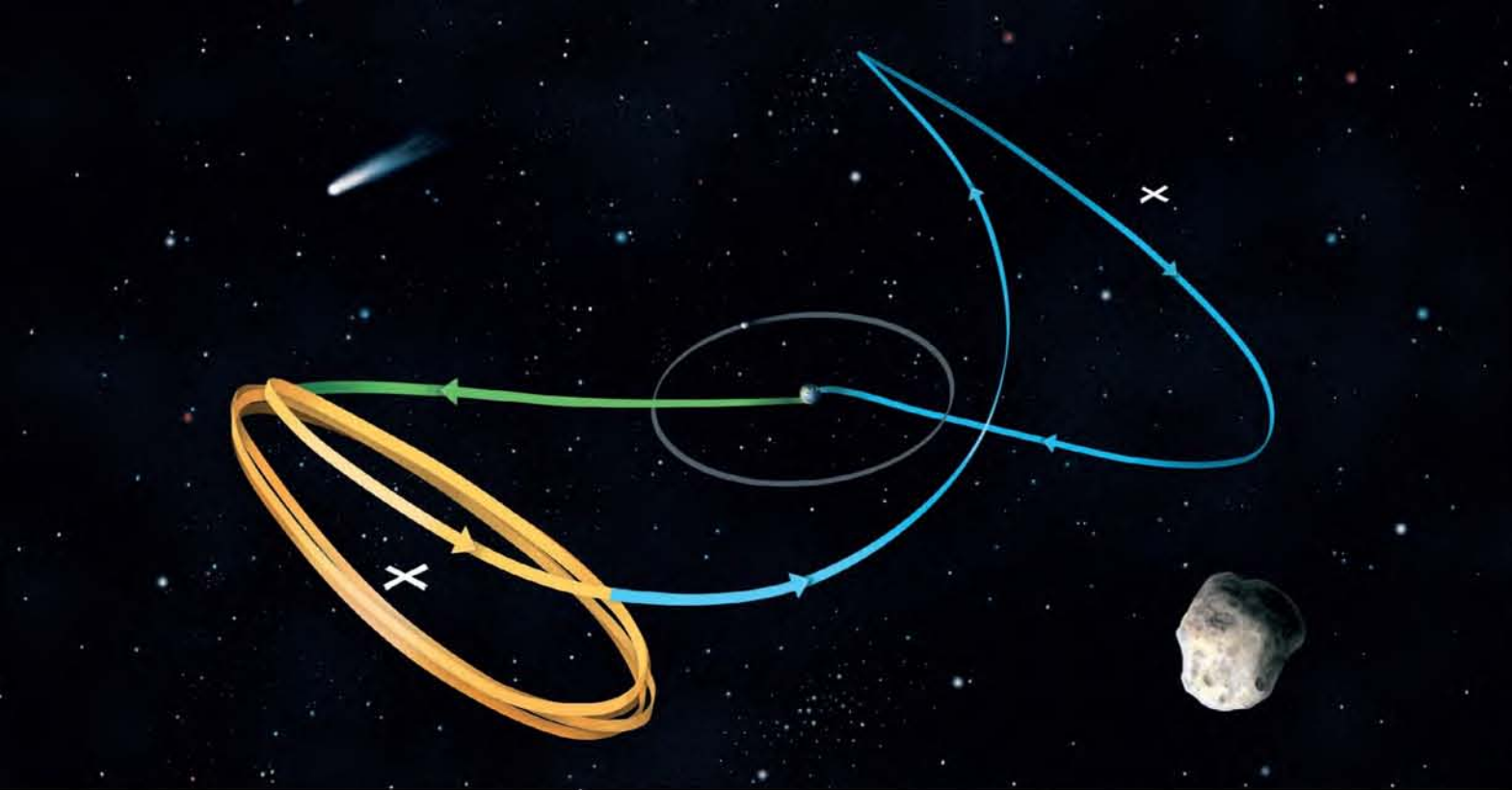
JPL/NASA

Spesso ci si raffigura il sistema solare come un meccanismo cosmico di precisione. Con poche eccezioni, i pianeti orbitano intorno al Sole seguendo traiettorie quasi circolari, i satelliti ruotano intorno ai loro pianeti, e tutti i movimenti di questi corpi obbediscono alla proverbiale immutabilità dei cieli. Si potrebbe immaginare che il campo gravitazionale generato da un meccanismo così ordinato sia egualmente regolare. Se si mette una pietra o un veicolo spaziale vicino al Sole, l'oggetto dovrebbe precipitare nell'immane massa della stella. Se lo si lascia andare vicino alla Terra, dovrebbe ricadere, un po' più lentamente, sul nostro pianeta.

Purtroppo la natura non è così semplice. La complicazione principale sta nel fatto che la Terra non è ferma in un punto dello spazio, ma orbita intorno al Sole. Così una pietra lasciata libera nello spazio nei pressi del nostro pianeta potrebbe trovarsi a seguire un cammino complesso e caotico, ruotando, per esempio, prima intorno alla Terra, poi intorno al Sole e di nuovo intorno alla Terra, continuando per anni. Se si aggiunge al quadro l'attrazione gravitazionale degli altri pianeti e satelliti, allora le possibili traiettorie possono diventare enormemente complicate e interessanti.

Di recente studiosi delle discipline più disparate, dalla matematica alla chimica alla dinamica dei fluidi, hanno dimostrato l'esistenza di un complesso insieme di traiettorie consentite agli oggetti in movimento: una rete interplanetaria di trasporto costituita da percorsi che si incrociano. Queste vie di comunicazione invisibili, che hanno origine presso un pianeta o un satellite, incanalano il traffico nel sistema solare. Ma, al contrario delle strade terrestri, quelle spaziali e i loro punti di scambio sono in continuo mutamento, e le corsie si spostano l'una rispetto all'altra a seconda delle mutevoli configurazioni di pianeti e satelliti. Procedendo a fatica in questa ragnatela, comete e asteroidi si trovano a balzare all'improvviso da una corsia all'altra, a essere attratti prima da un pianeta e poi da un altro, o talvolta a collidere con uno di essi. Per questi piccoli corpi cosmici, il sistema solare è più simile a un mare in tempesta che a un meccanismo a orologeria.

Anche se non si può influenzare la traiettoria di comete e asteroidi lanciati a tutta velocità (non ancora, per lo meno), i progettisti di missioni spaziali sono in grado di manovrare i loro veicoli e possono farli passare da una corsia a un'altra di un'autostrada interplanetaria in modo che percorrano enormi distanze con un consumo di combustibile quasi nullo.

Jen Christiansen

I VOLI SPAZIALI RICHIEDONO UNA NOTEVOLE QUANTITÀ DI CARBURANTE. Ma, dopo la partenza dalla Terra, la sonda Genesis ha percorso 1,5 milioni di chilometri in direzione del Sole (*in verde*), una distanza pari a circa quattro volte quella dell'orbita lunare (*in grigio*). Genesis è poi entrata in orbita intorno al punto lagrangiano terrestre L_1 (*crocetta bianca in primo piano*). In questa posizione ha raccolto per due anni e mezzo le particelle del vento solare, e quindi ha seguito una complicata traiettoria lunga milioni di chilometri (*in blu*), che girava intorno al secondo punto lagrangiano, L_2 (*l'altra crocetta bianca*), prima di far ritorno a Terra nel settembre 2004. La cosa stupefacente è che Genesis ha compiuto il suo viaggio quasi senza consumare combustibile. La sonda ha seguito una delle possibili traiettorie a basso dispendio di energia che attraversano il sistema solare e che fungono da condotti naturali tra i pianeti, percorsi da comete e asteroidi. Alcuni di questi condotti entrano in collisione con la Terra, come ha fatto, di proposito, la traiettoria di Genesis.

Proprio come un naufrago su un'isola deserta può affidare un messaggio in bottiglia alla corrente giusta nel momento giusto, i controllori possono dirigere un veicolo spaziale in particolari siti favorevoli dal punto di vista gravitazionale, che rappresentano porte d'accesso privilegiate verso destinazioni più remote. Sfruttando opportunamente questo effetto, è possibile progettare rotte efficienti sotto il profilo dei consumi, che in altro modo non sarebbero immaginabili o tecnicamente fattibili.

Come Buck Rogers

Nella maggior parte dei casi i voli interplanetari non sfruttano particolari effetti gravitazionali: i veicoli spaziali si dirigono rapidamente verso la loro destinazione nello stile di Buck Rogers, eroe di uno dei primi fumetti di fantascienza, sfruttando razzi a propulsione chimica. È un modo molto semplice di muoversi nello spazio. Chi pianifica la traiettoria deve considerare solo l'influenza di un corpo celeste per volta. Ciò significa che è possibile trattare sia la partenza dalla Terra sia l'arrivo su un pianeta lontano come interazioni separate tra il veicolo spaziale e un solo corpo. Allo stesso modo, il percorso di trasferimento dalle vicinanze della Terra alla zona circostante il pianeta di arrivo può essere elaborato considerando solo il veicolo spaziale e il Sole. In tutti i casi, quindi, si devono trattare solo due corpi per volta.

La traiettoria complessiva può allora essere approssimata per mezzo di una serie di curve semplici opportunamente connesse: ellissi, iperboli e parabole, ossia le cosiddette «soluzioni coniche» (sono infatti le curve generate dall'intersezione di un cono con un piano) al problema dei due corpi, scoperte da Keplero nel XVI secolo. Le spettacolari missioni della NASA con avvicinamenti ripetuti a pianeti del sistema solare esterno, come Voyager 1 e 2, seguivano un approccio basato sulle coniche, impiegato come prima approssimazione per una procedura numerica che generava soluzioni più precise tenendo conto di tutte le influenze, gravitazionali e non, sul veicolo spaziale.

Nel caso delle missioni che prevedono una successione di incontri con i pianeti – Giove, Saturno, Urano e Nettuno nel caso di Voyager 2 – la velocità del veicolo rispetto ai corpi planetari è elevata, e il periodo durante il quale il Sole e un pianeta possono produrre accelerazioni confrontabili sul veicolo stesso è molto breve. Per questo l'approccio basato sulle coniche funziona.

Gli inconvenienti, però, non mancano: gli incontri planetari sono fuggevoli, e il combustibile diventa un fattore limitante nella scelta dell'itinerario. Ce ne vorrebbe una quantità proibitiva, per esempio, per rallentare una sonda e inserirla in orbita intorno a un pianeta o a un satellite lontano, compiere osservazioni per un certo tempo e poi proseguire verso la destinazione successiva. Inoltre, caricare a bordo combustibile aggiuntivo per le manovre costringe a ridurre la strumentazione scientifica. Quindi i progettisti delle missioni devono raggiungere un compromesso tra la traiettoria che si propongono di seguire e la quantità di strumenti che si possono

NASA

I PROGETTISTI DI MISSIONI SPAZIALI sanno che in certi casi le traiettorie migliori non sono le più dirette. A volte può essere vantaggioso sfruttare i percorsi a basso dispendio di energia che collegano particolari punti dello spazio. Per esempio una sonda che deve raggiungere la Luna potrebbe arrivare a destinazione transitando per uno dei punti lagrangiani lunari, L_1 o L_2 (*in verde*). Gli stessi punti potrebbero essere stazioni di sosta nei voli verso altri pianeti, come indica questa mappa della «metropolitana spaziale».

Jimmy Dorff

L'ESISTENZA DI PERCORSI A BASSO DISPENDIO DI ENERGIA nello spazio può essere compresa considerando la fisica di un «pozzo gravitazionale». Questi dispositivi a forma di imbuto (come quello nella foto, che si trova al Morehead Planetarium di Chapel Hill, in Nord Carolina) permettono a una moneta di percorrere un'orbita stabile, più o meno nello stesso modo in cui un pianeta gira intorno al Sole.

In sintesi/*Le autostrade del cosmo*

- La dinamica del sistema solare non è così scontata come si potrebbe pensare. L'interazione di corpi di piccola massa con il Sole e i pianeti può produrre traiettorie singolari, percorsi a basso dispendio di energia lungo i quali una sonda spaziale potrebbe muoversi senza consumare carburante.
- L'approfondimento delle conoscenze sul «problema dei tre corpi» sarà certamente prezioso per la colonizzazione dello spazio, ma l'idea dei percorsi a basso consumo di energia ha applicazioni anche in altri campi, come la chimica e la fisica atomica.

trasportare. Nel caso della sonda Galileo, diretta su Giove, e del lander lunare delle missioni Apollo, per esempio, il combustibile caricato al momento del lancio era la metà della massa totale del veicolo.

In una categoria completamente diversa ricade la Genesis Discovery Mission per lo studio del vento solare, in cui il combustibile rappresentava appena il cinque per cento della massa. Lanciata nel 2001, Genesis ha percorso 1,5 milioni di chilometri in direzione del Sole, è rimasta per due anni e mezzo a raccogliere singoli atomi del vento solare e nel 2004 li ha riportati sulla Terra percorrendo complessivamente 32 milioni di chilometri.

Sfortunatamente il paracadute che doveva frenare la discesa della sonda non si è aperto, e il contenitore dei campioni si è gravemente danneggiato precipitando al suolo. Tuttavia è stato possibile recuperare una parte di ciò che era stato raccolto: i primi campioni di materia extraterrestre riportati a Terra dallo spazio dopo la fine del programma Apollo, e i primi in assoluto provenienti da oltre l'orbita lunare. Genesis ha compiuto il suo viaggio consumando una minima quantità di carburante. Per fare un confronto, pensiamo che un'automobile con il pieno di benzina (anch'essa circa il cinque per cento della massa totale del veicolo) può percorrere solo 500 chilometri prima di dover fare rifornimento.

Voli come quello della sonda Genesis sarebbero stati inconcepibili fino a poco tempo fa, ma oggi sono divenuti possibili grazie alla migliore conoscenza delle «autostrade interplanetarie». Concettualmente, l'approccio necessario per progettare un simile viaggio nello spazio non è molto diverso da ciò che i marinai fanno da tempo immemorabile: sfruttare le correnti per giungere velocemente a destinazione. Gli antichi navigatori hanno scoperto queste correnti osservando il moto di detriti e alghe trascinati dalle correnti stesse. Nella moderna navigazione spaziale in un certo senso si fa la stessa cosa osservando il moto di comete e asteroidi.

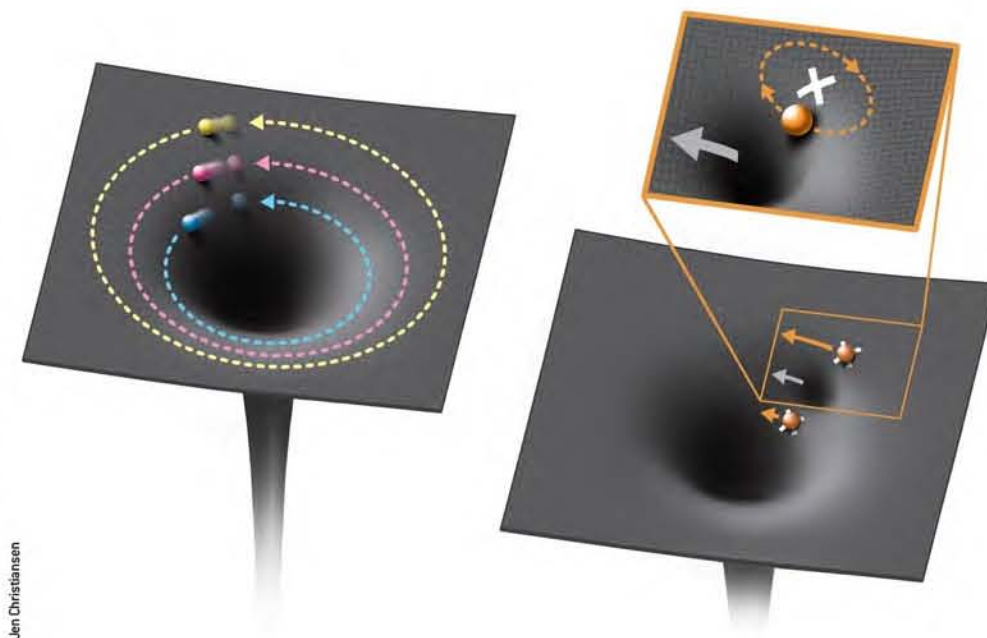
A questo proposito è particolarmente interessante la cometa Oterma. All'inizio del XX secolo questo corpo ghiacciato ruotava intorno al Sole su un'orbita esterna a quella di Giove. Poi, dopo essere transitata vicino al pianeta nel 1937, Oterma ha trasferito la propria orbita all'interno di quella di Giove. I due corpi si sono

nuovamente incontrati nel 1963, e a questo punto la cometa è tornata a orbitare esternamente a Giove, cosa che continua a fare ancora oggi. Durante ciascuno dei suoi incontri, ha seguito un'ampia orbita intorno al pianeta gigante, e ciò significa che per un certo tempo Oterma è stata una delle lune di Giove.

Perché questa cometa ha seguito una traiettoria così strana? Il modo migliore per capirlo è semplificare il problema e considerare il moto della cometa (un oggetto relativamente piccolo) come determinato dall'attrazione gravitazionale di due corpi: in questo caso, il Sole e Giove. Nello studio della meccanica celeste, questa situazione è denominata «problema dei tre corpi ristretto» (dal vincolo che il terzo corpo abbia massa trascurabile rispetto agli altri due). La soluzione integrale del problema è piuttosto difficile da elaborare, ma gli elementi chiave si possono comprendere ricorrendo a concetti fisici abbastanza intuitivi.

Tubi e imbuti

Basandosi sui lavori di Keplero, Isaac Newton risolse il problema gravitazionale dei due corpi. La formula che ne risulta è relativamente semplice, e può essere utilizzata per calcolare orbite ellittiche o incontri ravvicinati di veicoli spaziali su traiettorie iperboliche. Ma determinare il percorso seguito da una cometa o da una sonda spaziale che si trovi sotto l'influsso gravitazionale di due corpi orbitanti l'uno intorno all'altro è molto più difficile. Ipotizzare che uno dei corpi sia molto più piccolo degli altri due aiuta a rendere il problema più semplice, ma trovare la soluzione è comunque arduo. Sono molti gli scienziati che vi si sono cimentati, ma persino Newton dovette desistere. Voleva ricavare un'equazione che descrivesse il moto del terzo corpo in ogni istante: non ci riuscì, e il problema dei tre corpi fu dichiarato irrisolvibile.

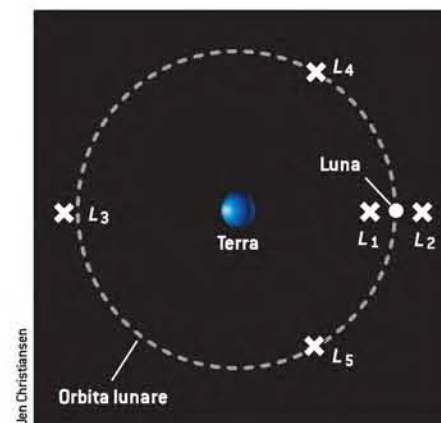
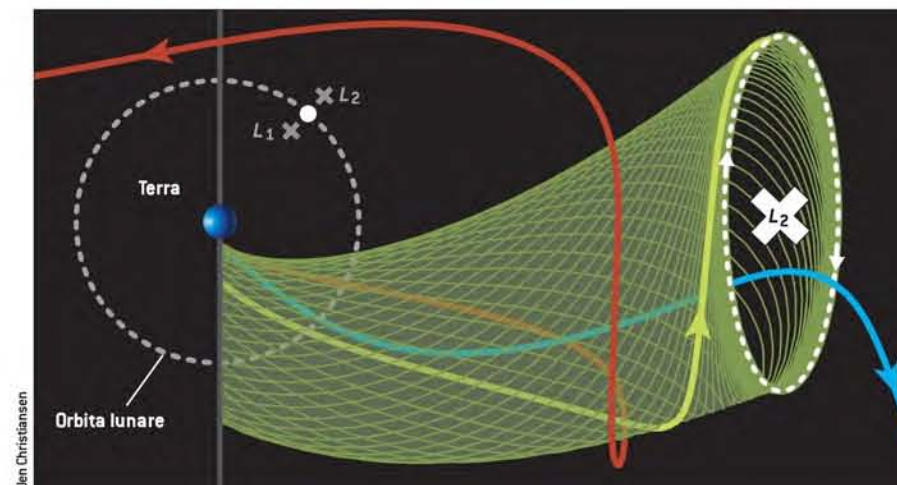
Ma, non arrendiamoci così facilmente. Il trucco non consiste nel cercare un'equazione che descrive il moto, ma nel considerare il problema da un punto di vista geometrico e cercare di cogliere intuitivamente come potrebbero presentarsi le soluzioni. A questo scopo, è utile pensare a uno di quei marchingegni a forma di imbuto che qualche volta si vedono nei musei scientifici e che spesso prendono il

LA VELOCITÀ DELLE BIGLIE CHE RUOTANO IN UN POZZO GRAVITAZIONALE – o dei pianeti che orbitano intorno al Sole – dipende dall'ampiezza dell'orbita: quanto più questa è piccola, tanto maggiore è la frequenza angolare necessaria per mantenere l'equilibrio (*a sinistra*). Il campo gravitazionale generato da due corpi come il Sole e la Terra ha la forma di un grande imbuto contenente un imbuto più piccolo (*in basso a destra*). Il piccolo imbuto deve orbitare intorno a quello maggiore, come fa la Terra con il Sole. Una biglia che ruoti intorno al grande imbuto con la stessa frequenza angolare può trovarsi in equilibrio in due punti ai lati opposti dell'imbuto piccolo (*crocette bianche*), corrispondenti ai punti lagrangiani terrestri L_1 e L_2. Con molta attenzione, si potrebbe mettere una biglia in posizione appropriata e conferirle una velocità iniziale tale da farla entrare in «orbita» intorno a questi punti speciali, proprio come la missione Genesis è stata progettata per orbitare intorno al punto lagrangiano terrestre L_1 (*in alto a destra*).

L'AUTORE

SHANE D. ROSS, è un *postdoctoral fellow* in scienze matematiche alla University of Southern California, da dove sta per trasferirsi al Virginia Polytechnic Institute and State University. I suoi interessi di ricerca includono lo studio del controllo dei veicoli spaziali, i metodi geometrici per sistemi ingegneristici e i processi di mescolamento e trasporto. Questo articolo è stato pubblicato sul numero di maggio-giugno 2006 di «American Scientist».

nome di «pozzo gravitazionale», a causa dell'analogia fisica con la gravitazione. Un piccolo scivolo su un lato fa cadere una moneta nello strumento in modo che essa continui a ruotare all'interno dell'imbuto per un tempo notevolmente lungo. Una biglia si muoverebbe allo stesso modo.

Se non ci fosse l'attrito, la moneta continuerebbe a muoversi in cerchio indefinitamente, imitando l'orbita della Terra intorno al Sole (in questo caso l'imbuto rappresenterebbe il pozzo gravitazionale solare). Ma nel mondo reale la moneta perde un po' di energia a causa dell'attrito e, col tempo, comincia a muoversi a spirale verso l'interno e verso il basso. Questa tendenza è facile da osservare e, in effetti, è la motivazione economica per l'installazione di questi strumenti: alla fine si mangiano la moneta. Ma guardandola scomparire si impara qualcosa. In particolare si nota che quanto più l'orbita è piccola tanto maggiore è il numero di giri che la moneta compie in un dato tempo. Ciò significa che un'orbita piccola ha una frequenza angolare elevata, e la stessa cosa vale per gli oggetti in orbita nello spazio.

Immaginiamo ora un imbuto con tre biglie che ruotano seguendo altrettante orbite parallele separate da una distanza molto breve. Rispetto alla biglia centrale, quella nell'orbita esterna deve muoversi a una frequenza angolare inferiore per essere stabile, mentre quella nell'orbita interna deve spostarsi a una frequenza angolare più alta. La stessa cosa vale nello spazio, per esempio nel caso di tre asteroidi in orbita intorno al Sole. Se l'orbita

NEL PROBLEMA DEI TRE CORPI RISTRETTO esistono cinque punti di equilibrio gravitazionale, qui mostrati per il sistema Terra-Luna. Tutti si spostano con la Luna, via via che essa orbita intorno alla Terra. I punti lagrangiani L_1, L_2 e L_3 sono considerati instabili, perché un oggetto collocato in uno di essi tende ad allontanarsi lentamente col tempo. I punti lagrangiani L_4 e L_5 sono stabili, nel senso che gli oggetti posti nelle loro vicinanze tendono naturalmente a rimanervi.

UN VEICOLO SPAZIALE a cui è data un'opportuna velocità iniziale può essere inserito in una traiettoria tale da portarlo in orbita, per esempio, intorno al punto lagrangiano terrestre L_2 (*curva in verde chiaro*). Una collezione di simili traiettorie (ogni punto nel quale la sonda avrebbe una specifica posizione e velocità) costituisce un «tubo» della rete di trasporto interplanetaria (*reticolo verde*). Una sonda che si trovi su una traiettoria nell'interno del tubo oltrepassa L_2 e si dirige verso il sistema solare esterno (*curva in blu*), mentre un veicolo che si trovi su una traiettoria esterna al tubo ritorna verso il Sole (*curva in rosso*).

del corpo centrale fosse posta a un'unità astronomica di distanza dalla stella (ossia la distanza Terra-Sole, circa 150 milioni di chilometri), impiegherebbe 365 giorni per compiere una rivoluzione. All'asteroide esterno occorrerebbe qualche giorno in più, mentre quello interno impiegherebbe qualche giorno in meno.

Tutto ciò è facile da capire se si ricorda ciò che accade alla moneta nell'imbuto. Ma supponiamo che nel nostro marchingegno ci sia un piccolo imbuto inserito in quello principale, e supponiamo anche che esso ruoti dentro l'imbuto maggiore alla stessa velocità di una moneta o di una biglia in quella posizione. Una configurazione simile riproduce la combinazione del pozzo gravitazionale terrestre (l'imbuto piccolo) e solare (quello più grande).

Consideriamo ora ciò che accadrebbe se si lanciasse nel dispositivo una biglia in modo da farla ruotare alla stessa distanza dal centro e alla stessa velocità dell'imbuto piccolo (la Terra), ma non troppo vicino a esso. La biglia continuerebbe a girare indisturbata per parecchio tempo. Se fosse in moto a questa stessa velocità angolare, ma in posizione più esterna (dove la superficie ha una pendenza moderata), la biglia sarebbe troppo veloce per mantenere un'orbita stabile e verrebbe scagliata fuori dal dispositivo. Viceversa, se fosse più all'interno rispetto all'imbuto che rappresenta la Terra, sarebbe troppo lenta per mantenere la propria quota sulla ripida parete del dispositivo e cadrebbe verso il basso. L'unica posizione in cui una biglia con quella particolare frequenza angolare può essere stabile corrisponde al raggio dell'orbita dell'imbuto piccolo.

Un esame più attento di questa strana superficie rivela l'esistenza di due punti molto particolari. Uno si trova vicino alla cresta che collega il piccolo imbuto terrestre a quello principale. Arriviamoci partendo dal centro del dispositivo (nelle profondità dell'imbuto solare) e spostandoci in direzione dell'imbuto terrestre. La superficie dapprima si alza con la consueta ripidità, ma poi la pendenza diminuisce fino ad annullarsi, e quindi si ricade nell'imbuto terrestre. Prima di arrivare alla sommità della cresta intermedia si passa in un punto dove la pendenza è perfetta per mantenere in equilibrio una biglia che ruoti alla stessa velocità dell'imbuto minore. Ricordiamo che, in condizioni normali, questa biglia dovrebbe muoversi più velocemente per rimanere in equilibrio. Ma, poiché qui la parete ha una pendenza un po' meno ripida di quella che tipicamente si dovrebbe incontrare in corrispondenza di questo raggio orbitale, la biglia può seguire un'orbita stabile. Questa posizione di equilibrio per la biglia ha un'analoga nello spazio: si trova a 1,5 milioni di chilometri dalla Terra, in direzione del Sole.

Tornando al nostro dispositivo, all'esterno dell'imbuto che rappresenta la Terra c'è un altro punto speciale di equilibrio. Ricordiamo che, in questa posizione, una biglia orbiterebbe a una frequenza angolare inferiore a quella dell'imbuto terrestre. Se avesse la stessa frequenza angolare, ma si trovasse su un'orbita più esterna, dove la pendenza dell'imbuto principale è più dolce, normalmente dovrebbe essere scagliata via. Vi è però un punto – appena all'esterno dell'imbuto terrestre, dove la superficie è un po' più ripida del normale – dove questo non accade. Anche in questo caso esiste un punto equivalente nello spazio, situato a 1,5 milioni di chilometri dalla Terra, in direzione opposta al Sole.

Eulero, celebre matematico svizzero del XVIII secolo, scoprì questi due punti speciali, e anche un terzo. Il suo contemporaneo Lagrange ne individuò altri due, e oggi tutti e cinque sono detti punti lagrangiani. Anche se ciascuno di essi rappresenta una particolare orbita intorno al Sole, si chiamano «punti» perché appaiono come posizioni fisse in un sistema di riferimento che ruoti alla stessa velocità con cui la Terra e il Sole orbitano intorno al loro comune cen-

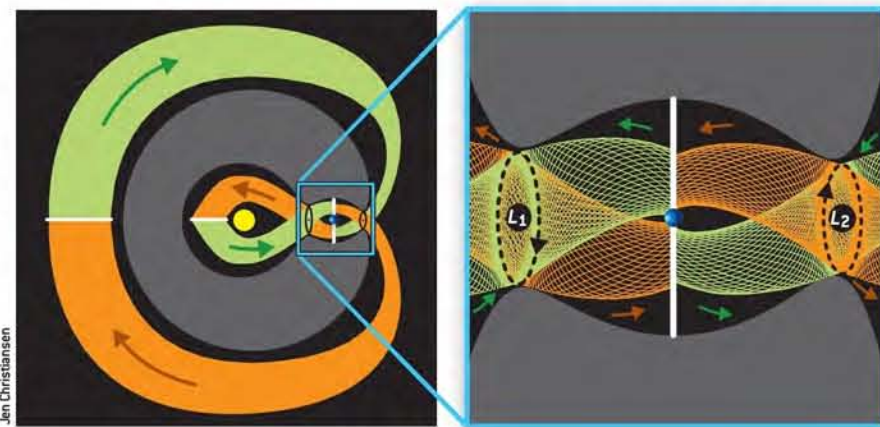
tro di massa (che si trova in profondità nel Sole). Cinque punti speciali del tutto analoghi, indicati con lettere che vanno da L_1 a L_5, esistono per ogni coppia di corpi dotati di massa: il Sole e un pianeta, un pianeta e uno dei suoi satelliti e così via. L_1 corrisponde al punto di equilibrio più interno per la biglia descritto in precedenza, mentre L_2 corrisponde al punto di equilibrio più esterno. Entrambi sono interessanti per comprendere la rete interplanetaria di trasporto, dato che sono portali privilegiati per raggiungere destinazioni lontane.

L_1 e L_2 sono classificati come punti di equilibrio instabile, ma questa denominazione può essere fuorviante, perché un veicolo spaziale può rimanervi a lungo. Anzi, le lievi interazioni tra forze gravitazionali e rotazionali permettono al veicolo di «orbitare» intorno a L_1 o L_2 nel sistema di riferimento in rotazione anche se nei due punti non vi è alcun oggetto materiale. Per quanto simili orbite intorno a un punto dello spazio possano apparire bizzarre, rappresentano l'approssimazione reale rispetto al trovarsi esattamente in L_1 o L_2 e in moto alla velocità esatta per conseguire una stabilità perfetta.

Per capire meglio, immaginiamo che una sonda spaziale si trovi in orbita intorno al Sole vicino al punto L_2 terrestre, ma lievemente all'interno rispetto a esso. Supponiamo anche che sia in moto a una velocità di poco superiore a quella che dovrebbe avere se si trovasse esattamente in L_2. Che cosa accade? Di nuovo, è utile visualizzare una biglia che ruota entro un dispositivo a doppio imbuto. Se avesse posizione e velocità corrispondenti a quelle del veicolo spaziale, la biglia comincerebbe a spostarsi in avanti (in relazione a L_2) nell'orbita che descrive intorno all'imbuto rappresentante il Sole; tenderebbe anche a essere sospinta leggermente verso l'esterno.

Ma la superficie in questa zona ha una forma peculiare. Mentre la biglia si sposta in avanti e verso l'esterno, la superficie di fronte a essa si innalza e la costringe a rallentare, e ben presto L_2 la raggiunge (stando più all'interno). La biglia allora comincia a seguire L_2 e trova nuovamente la superficie che si innalza dietro di sé, accelerandola e dirigendola verso il Sole, proprio come un'onda sospinge una persona che fa *surf* verso la riva (e spesso un poco lateralmente). In definitiva, la biglia torna più o meno nel punto di partenza e all'incirca con la stessa velocità, e potrebbe riuscire a «orbitare» in questo modo per alcune volte intorno a L_2 prima di essere scagliata verso l'esterno o di cadere nel vicino imbuto che rappresenta la Terra.

Un veicolo spaziale situato nei pressi di L_1 può comportarsi in maniera analoga. Visto da un osservatore a terra, sembrerebbe orbitare per un certo tempo intorno al punto lagrangiano terrestre L_1 e poi partire velocemente verso la Terra o cominciare a ruotare intorno al Sole, e il tutto senza consumare combustibile. Alcune delle possibili orbite intorno ai punti lagrangiani giacciono nel piano dell'orbita terrestre. Altre, come quella seguita dalla sonda Genesis, sono tridimensionali, con una varietà di forme spiraleggianti, ed entrano ed escono dal piano orbitale dei due corpi.

Jen Christiansen

ALCUNI TUBI DELLA RETE INTERPLANETARIA di trasporto portano un oggetto a inserirsi in orbita intorno ai punti lagrangiani terrestri L_1 e L_2 (*traiettorie in verde, a destra*), mentre altri li allontanano da queste orbite (*in arancione*). I progettisti di missioni spaziali possono sfruttare le intersezioni di questi tubi di ingresso e di uscita, facendo in modo che una sonda salti da uno all'altro spostandosi tra L_1 e L_2 o inserendosi in orbite circumsolari che possono essere più piccole o più grandi di quella della Terra (*a sinistra*). Tuttavia, un veicolo con energia cinetica limitata (per esempio sufficiente solo per orbitare intorno a L_1 o L_2) non può visitare certe regioni (*in grigio*), indipendentemente dalla traiettoria seguita.

Surf interplanetario

Alla fine del XIX secolo il matematico francese Henri Poincaré compì significativi passi avanti nell'interpretazione della meccanica celeste relativa a questo caso. Fu il primo a rendersi conto di quanto potesse essere complicato il moto del terzo corpo. I metodi geometrici che impiegò per giungere a questa conclusione hanno rappresentato le basi di quella disciplina che oggi prende il nome di dinamica non lineare, o più generalmente di teoria del caos.

È importante avere chiaro che «caotico» non significa casuale. In questo problema esistono traiettorie che appaiono caotiche, ma che in realtà sono prevedibili, almeno per un certo intervallo di tempo. Perciò è possibile sfruttarle per progettare diverse rotte nello spazio che consentono un basso consumo di energia. Poincaré portò ordine nel caos organizzando queste traiettorie in speciali collezioni di superfici che esistono in quello che i matematici chiamano «spazio delle fasi esadimensionale», in cui troviamo le tre dimensioni dello spazio normale (diciamo x, y e z) e tre dimensioni che rappresentano la velocità di un oggetto in ciascuna direzione.

Partendo dal lavoro di Poincaré, alla fine degli anni sessanta Charles C. Conley, un matematico dell'Università del Wisconsin, scoprì una collezione di superfici a forma di tubo per oggetti posti sotto l'influenza gravitazionale di due corpi mutuamente orbitanti; questo risultato fu poi studiato a fondo da Robert P. McGehee, allora studente di Conley e oggi all'Università del Minnesota. Un oggetto situato su uno di questi tubi esadimensionali (vale a dire che abbia esattamente la posizione e la velocità associate a esso) viene naturalmente trasportato verso una traiettoria che descrive un'orbita intorno al punto lagrangiano L_1 o L_2, visto nel sistema di riferimento in rotazione, oppure viene allontanato da essa. Le traiettorie all'interno del tubo oltrepassano il punto lagrangiano, mentre quelle all'esterno finiscono per far rimbalzare l'oggetto all'indietro.

Jen Christiansen

LA COMETA OTERMA ha sfruttato la rete di trasporto interplanetaria passando da un'orbita esterna a quella di Giove (che percorreva nel 1910) a un'orbita interna a quella del pianeta, sulla quale è rimasta dal 1937 al 1963, per poi tornare su un'orbita simile a quella originaria. Il suo bizzarro vagabondaggio nello spazio è mostrato qui in un sistema di riferimento fisso (*a sinistra*) e in uno rotante (*a destra*).

L'ESPLORAZIONE DEI SATELLITI GHIACCIATI DI GIOVE potrebbe essere facilitata da una traiettoria accortamente progettata. Una sonda potrebbe, per esempio, entrare nel sistema gioviano lungo un tubo di ingresso (*fascia verde esterna a sinistra*), raggiungendo Ganimede e orbitando brevemente intorno a esso prima di seguire un tubo di uscita (*in arancione*) lungo il quale si sposterebbe in un'orbita intorno a Giove più piccola di quella di Ganimede. La sonda salterebbe poi in un tubo d'ingresso diretto verso il satellite Europa (*fascia verde interna*), e orbiterebbe intorno a quest'ultimo per un tempo significativo (*a destra*).

Traiettoria della sonda
Europa
L2

Jen Christiansen

A partire dalla metà degli anni novanta, ho collaborato con Martin W. Lo del Jet Propulsion Laboratory (JPL) della NASA e con Wang Sang Koon e Jerrold E. Marsden del California Institute of Technology per perfezionare questo approccio. Abbiamo dimostrato che un'importante proprietà fisica dei tubi è che qualsiasi oggetto che si sposti da un'orbita interna a quella di un pianeta a un'orbita esterna deve necessariamente transitare lungo i tubi. L'insieme dei possibili oggetti che scavalcano il pianeta «scorre» lungo questi tubi, come l'acqua che scorre in una canna da giardino, ma in sei dimensioni anziché in tre.

Il confronto con un liquido non è solo un'analogia. In effetti, strumenti informatici inizialmente progettati da François Lekien della Princeton University e collaboratori per simulare «canali» dinamici negli oceani sono stati applicati anche per la determinazione di traiettorie a basso consumo di energia nello spazio.

Calcolando la configurazione dei tubi meglio di quanto riuscirono a fare Conley e McGehee, i miei colleghi e io abbiamo scoperto che essi si estendono fino a grandi distanze dalla loro regione d'origine (vicino a L_1 o L_2) e si avvolgono intorno a una qualsiasi coppia di oggetti massicci presi in considerazione nel calcolo, allungandosi e torcendosi lungo il cammino. Si può pensare che esista una «regione porta» intorno a L_1 e un'altra intorno a L_2 e che i tubi rappresentino le vie d'accesso e di uscita della zona «controllata» da un dato pianeta o satellite. Un'altra proprietà degli oggetti che percorrono uno di questi tubi è che si muovono alla minima velocità relativa al vicino pianeta o satellite quando si trovano nella regione porta, che può essere immaginata come una zona di quasi equilibrio, il culmine di un rilievo di energia che gli oggetti devono risalire e superare.

Si vede così che la strana traiettoria di Oterma giace proprio lungo i percorsi tubolari connessi ai punti lagrangiani L_1 e L_2 di Giove: è quasi come se la cometa avesse seguito la galleria di una metropolitana interplanetaria che unisce regioni di spazio situate a grande distanza. Una parte di questi percorsi può addirittura attraversare il pianeta. È possibile che la cometa Shoemaker-Levy 9 stesse percorrendo proprio un tubo di questo genere nel 1994, quando si frammentò precipitando su Giove.

Guida per autostoppisti

Se fosse possibile farsi dare un passaggio da Oterma o da uno degli altri oggetti che seguono i percorsi tubolari tra i pianeti, si potrebbe girare gratis tutto il sistema solare. Ma perché aspettare che passi un asteroide diretto dove vogliamo? È sufficiente dirigere il proprio veicolo spaziale in uno di questi condotti celesti. Percorrere queste traiettorie preferenziali può ridurre radicalmente la quantità di combustibile necessaria per esplorare il sistema solare. E il posto giusto per cercare di cogliere una simile opportunità è molto vicino alla Terra.

Dato che ha una visuale libera sul cosmo, il punto lagrangiano terrestre L_2 è adatto a telescopi per osservazioni dello spazio profondo, mentre la regione intorno a L_1, che consente un'ottima visione del Sole, è ideale per strumenti destinati allo studio della nostra stella. In effetti, uno dei motivi per cui è stato possibile realizzare la missione Genesis era la sua particolare «orbita ad alone» intorno al punto lagrangiano terrestre L_1. Vista da un punto di osservazione sulla Terra, Genesis sembrava muoversi in un alone tutto intorno al Sole. Le orbite ad alone furono così

chiamate, inizialmente nel caso lunare, da colui che le scoprì negli anni sessanta, Robert Farquhar dell'Applied Physics Laboratory della Johns Hopkins University, al quale spetta in gran parte il merito di aver organizzato la prima missione che sfruttasse i punti lagrangiani, l'International Sun-Earth Explorer 3.

Genesis ha percorso una traiettoria a basso consumo di energia per raggiungere la sua orbita ad alone, l'ha seguita durante il periodo di raccolta dei campioni e poi è ritornata verso la Terra lungo un'altra traiettoria a basso consumo che compiva un anello presso L_2. Sfruttando la loro conoscenza dei tubi, il principale progettista della missione, Lo, insieme con Kathleen C. Howell della Purdue University e Brian Barden (all'epoca studente di Howell), trovarono il modo per consentire a Genesis di seguire questa peculiare traiettoria con un ridottissimo consumo di carburante. Fu un'impresa che suscitò grande interesse sia tra coloro che si occupano di astronautica sia tra i matematici.

In particolare, il lavoro compiuto nella missione Genesis ha ispirato Lo e il sottoscritto a esaminare più in dettaglio la dinamica della regione di spazio prossima alla Terra. Ci siamo resi conto che i punti lagrangiani L_1 e L_2, sia nel sistema Terra-Sole sia in quello Terra-Luna, sono svincoli e destinazioni importanti. Fortunatamente, i tubi che uniscono le zone circostanti a questi quattro punti lagrangiani sono configurati in modo da intersecarsi qua e là. All'incirca una volta ogni mese lunare, le orbite ad alone intorno ai punti lagrangiani lunari L_1 e L_2 si connettono a orbite ad alone intorno ai corrispondenti punti terrestri mediante traiettorie a consumo di energia molto basso o addirittura nullo. Le implicazioni di questa fortuita configurazione per l'esplorazione e la colonizzazione del sistema solare sono straordinarie.

Lo e io, con alcuni colleghi della NASA, abbiamo avanzato la proposta di costruire una stazione spaziale permanente nel punto lagrangiano lunare L_1, da usare come nodo di interscambio in grado di ampliare considerevolmente, ben oltre l'orbita terrestre bassa, la portata delle attività di esplorazione spaziale. La stazione sarebbe il punto di sosta più vicino lungo l'autostrada interplanetaria. Da qui un carico potrebbe essere spedito con trasporti merci, lenti ma dai consumi ridotti, mentre gli astronauti viaggerebbero su veicoli più veloci. Una navetta in partenza dalla stazione potrebbe raggiungere in poche ore qualsiasi punto della superficie lunare: una soluzione perfetta per il transito di passeggeri diretti sulla Luna.

Questo portale sarebbe anche un eccellente punto di partenza e di arrivo per voli interplanetari diretti verso Marte, gli asteroidi e il sistema solare esterno. Ci sono traiettorie naturali che consentono di viaggiare da un pianeta a un altro senza consumare combustibile, ma richiedono migliaia di anni per arrivare a destinazione. Solo gli asteroidi, le comete e le meteoriti marziane (frammenti di roccia espulsi da Marte in seguito a una collisione, che dopo un lungo peregrinare sono precipitati sulla Terra) hanno la pazienza necessaria.

I futuri telescopi spaziali destinati a essere messi in orbita presso i punti lagrangiani terrestri L_1 e L_2 potrebbero essere assemblati in questa stazione e trasferiti nelle loro posizioni definitive usando pochissimo combustibile. E quando gli strumenti avranno bisogno di manutenzione potranno essere altrettanto economicamente riportati nei pressi della stazione.

ACS Science & Engineering Team e NASA

UNA SCIA DI STELLE esce dalla galassia «Girino», UGC 10214, lungo un canale tubolare che si allunga per circa 280.000 anni luce. Questo condotto (l'equivalente galattico dei tubi che costituiscono la rete interplanetaria di trasporto) è stato generato dall'interazione gravitazionale con una galassia compatta che ora si vede spuntare dietro uno dei bracci di spirale del Girino.

Si potrebbero esplorare le lune di Giove **senza consumo di combustibile**

Ma per sfruttare i percorsi a basso dispendio energetico non è necessario restarsene in prossimità della Terra. Io faccio parte di un gruppo internazionale (che comprende Koon, Marsden, Lo, Gerard Gómez, dell'Università di Barcellona, e Josep Masdemont, del Politecnico della Catalogna, sempre a Barcellona) che ha proposto una nuova classe di missioni spaziali. La nostra idea è che una sola sonda possa entrare successivamente in orbita intorno a diversi satelliti di uno dei pianeti esterni per effettuare osservazioni di lunga durata.

Per esempio, un *orbiter* potrebbe esplorare l'uno dopo l'altro Callisto, Ganimede ed Europa – satelliti di Giove di dimensioni planetarie, sui cui probabilmente c'è dell'acqua – seguendo una traiettoria che richiede una quantità di combustibile non proibitiva dal punto di vista tecnologico. La NASA ha preso in considerazione un progetto di questo genere, chiamato Jupiter Icy Moon Orbiter, che sfrutterebbe i collegamenti tra i tubi di bassa energia di Giove e dei suoi satelliti, ma i finanziamenti per la missione sono stati tagliati lo scorso anno e le sue prospettive sono incerte.

Dagli atomi alle galassie

L'approfondimento delle conoscenze sul problema dei tre corpi ristretto e sulla dinamica associata ai punti lagrangiani sarà certamente prezioso per l'esplorazione e la colonizzazione dello spazio. Ma si è visto che l'idea dei percorsi a basso consumo di energia ha applicazioni più vaste.

Il merito della prima intuizione spetta a Charles Jaffé, chimico della West Virginia University, il quale nel 2000 osservò che, in opportune condizioni sperimentali, le traiettorie degli elettroni di valenza nei cosiddetti atomi di Rydberg (particolari atomi i cui elettroni di valenza percorrono traiettorie lontane dal nucleo atomico ionizzato) assomigliano molto a quella seguita dalla sonda Genesis. E in effetti si è rilevato che, quando sono soggetti a campi elettrici e magnetici perpendicolari tra loro, anche gli atomi di Rydberg percorrono cammini tubolari. Jaffé ha iniziato a collaborare con me, Marsden, Lo, Turgay Uzer del Georgia Institute of Technology e David Farrelly della Utah State University per applicare le tecniche della chimica statistica allo studio del comportamento della materia marziana espulsa nello spazio a seguito di un impatto. Questo lavoro rappresenta la prima applicazione alla meccanica celeste di una tecnica ormai acquisita in chimica.

Questi scambi tra discipline diverse si sono sviluppati anche in altre direzioni. In collaborazione con Koon, Marsden, Tomohiro Yanao del Caltech, Frederic Gabern dell'Università di Barcellona e con un gruppo diretto da Michael Dellnitz dell'Università di Paderborn, in Germania, e da Oliver Junge del Politecnico di Monaco di Baviera, mi sono occupato dello sviluppo delle basi matematiche e computazionali di una teoria delle velocità di reazione che sia in grado di superare alcune tra le classiche difficoltà che si incontrano in chimica. Questo lavoro si è ispirato al calcolo del trasporto lungo tubi nel sistema solare e alle relative tecniche geometriche. Ovviamente sono gli aspetti matematici a stabilire un legame tra la chimica e la dinamica dei sistemi planetari.

È accertato che i tubi controllano struttura e moto anche a scala galattica: Toshi Fukushige dell'Università di Tokyo e Douglas Heggie dell'Università di Edimburgo hanno infatti dimostrato che tubi correlati ai punti lagrangiani determinano l'«evaporazione» di piccoli ammassi stellari in orbita intorno ad alcune galassie.

Esempi ancora più vistosi di strutture a tubo si osservano nelle galassie fortemente interagenti. Situata a 420 milioni di anni luce di distanza, UGC 10214, meglio nota come la galassia «Girino», evidenzia i segni di un breve ma violento episodio avvenuto in passato. La coda che si protende dal corpo della galassia indica la regione dove le stelle sono entrate in tubi che la collegano a una galassia intrusa che nel frattempo si è spostata e ora è in gran parte nascosta alla vista. La coda del Girino è quindi un ponte di 280.000 anni luce lanciato nel vuoto, ma anche altre coppie di galassie presentano condotti tubolari che le collegano.

Anche se non sono stati ancora individuati, si può presumere che tubi di questo tipo uniscano il sistema solare alle stelle vicine. Provate a immaginare che una delle due sonde Voyager, che ora hanno lasciato il sistema solare, sia entrata in un tubo diretto verso una regione di equilibrio di forze tra il Sole e, per esempio, Alfa Centauri, distante diversi anni luce: la sonda potrebbe ottenere un passaggio gratuito fino a un'altra stella. Anche se questo accadesse, data la loro velocità attuale, i Voyager impiegherebbero migliaia di anni per giungere a destinazione. Tuttavia nel lontano passato altre stelle si sono avvicinate al Sole molto più di quanto non facciano le nostre attuali vicine più prossime. È probabile che siano avvenuti scambi di materia tra il sistema solare e i sistemi stellari di passaggio, e che i tubi siano stati i canali invisibili attraverso cui sono avvenuti questi scambi.

I fan di Douglas Adams, autore della *Guida galattica per autostoppisti*, possono essere ottimisti: anche se richiederà tempi molto lunghi, fare l'autostop nella galassia potrebbe essere una realtà.

PER APPROFONDIRE

CONLEY C.C., *Low Energy Transit Orbits in the Restricted Three-Body Problem*, in «SIAM Journal of Applied Mathematics», Vol. 16, pp. 732-746, 1968.

FUKUSHIGE T. e HEGGIE D.C., *The Time-Scale of Escape from Stellar Clusters*, in «Monthly Notices of the Royal Astronomical Society», Vol. 318, pp. 753-761, 2000.

JAFFÉ C., ROSS S., LO M., MARSDEN J., FARRELLY D. e UZER T., *Statistical Theory of Asteroid Escape Rates*, in «Physical Review Letters», 89, 011101, 2002.

SMITH D.L., *Next Exit 0.5 Million Kilometers*, in «Engineering and Science», Vol. LXV, n. 4, pp. 6-15, 2002.

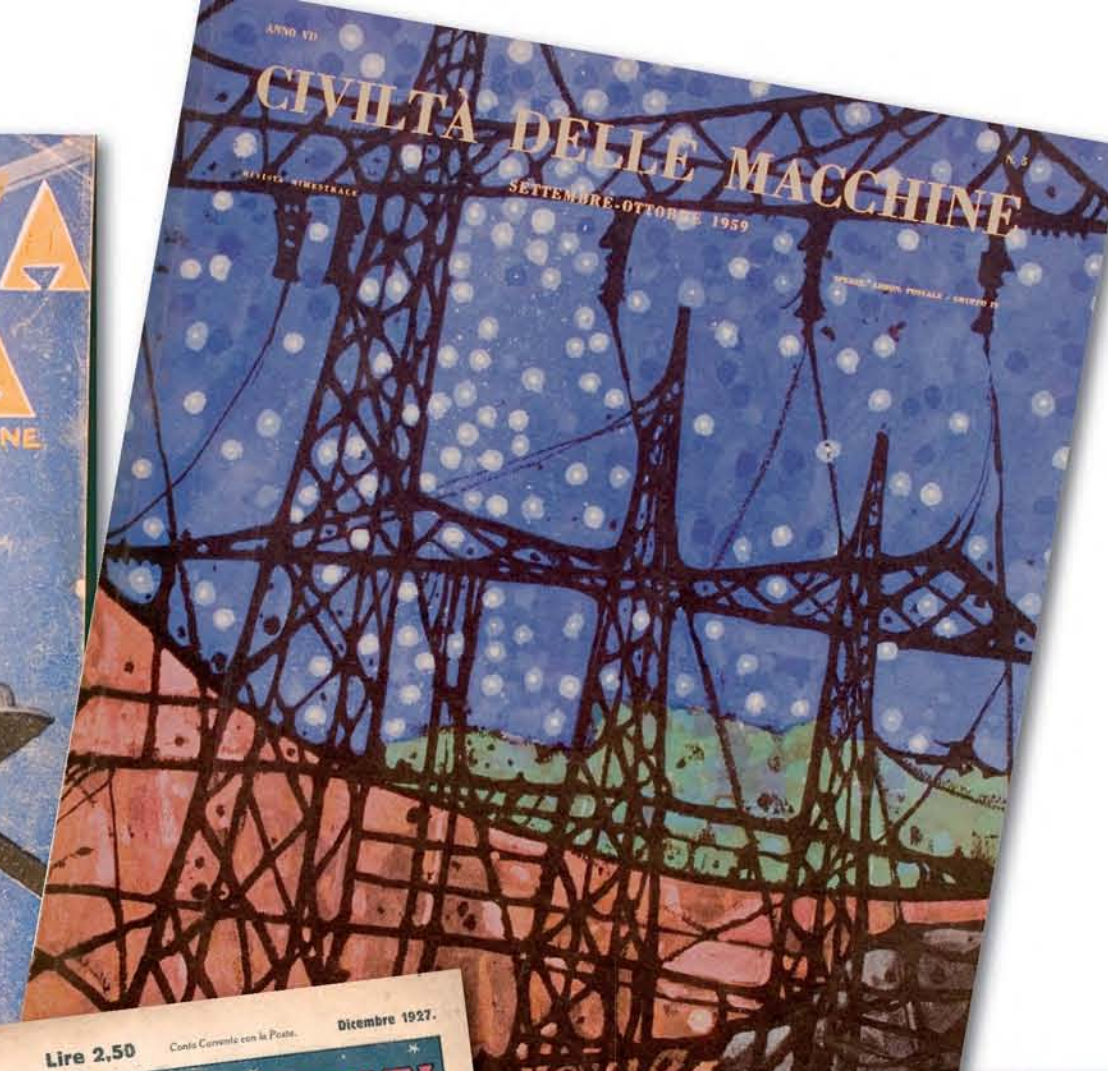
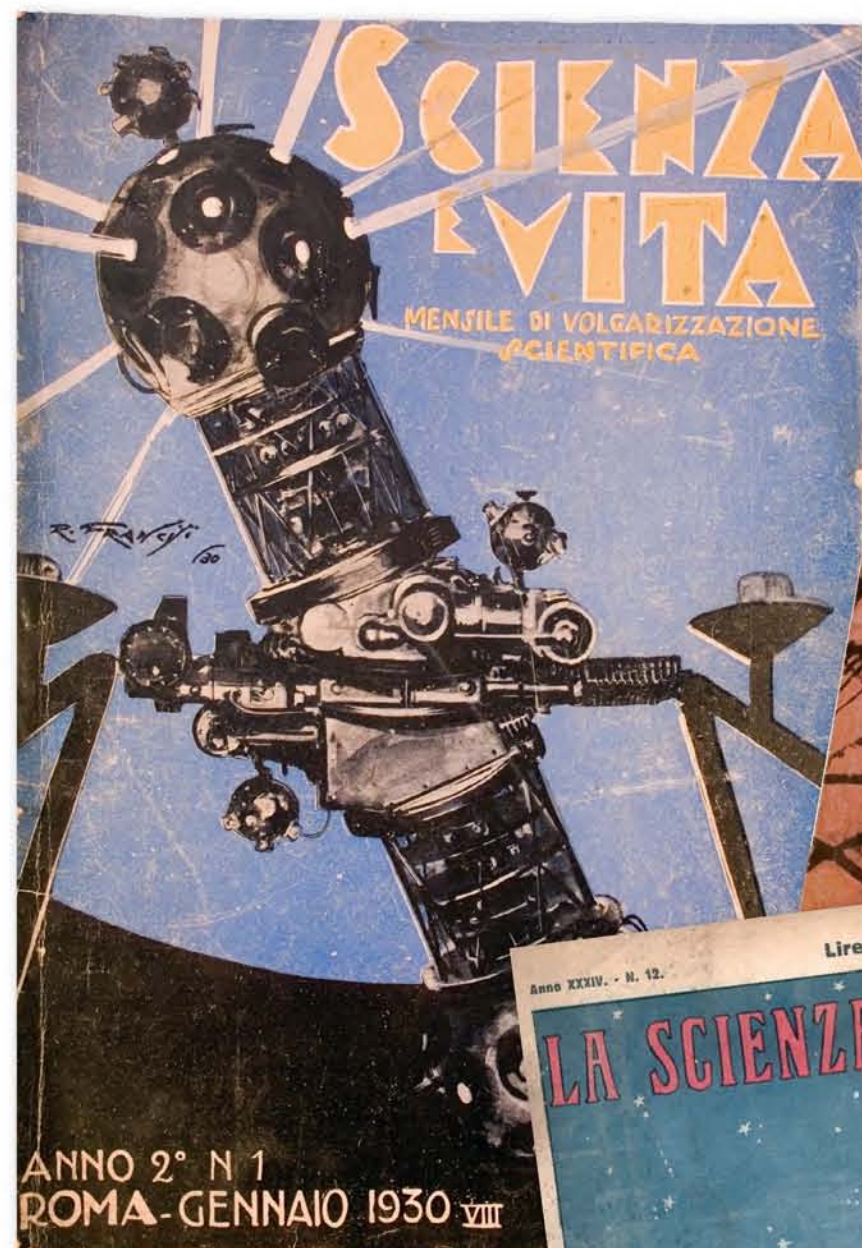
MARSDEN J.E. e ROSS S.D., *New Methods in Celestial Mechanics and Mission Design*, in «Bulletin of the American Mathematical Society», Vol. 43, pp. 43-73, 2006.

La scienza per tutti

Le riviste di divulgazione scientifica italiane hanno una storia antica e complessa. Da cui è possibile ricavare alcune lezioni molto attuali

di Enrica Battifoglia

La scienza piace. A volte diverte, perfino. Ci sono ambiti che attirano di più, come la medicina e la salute, e altri che da sempre affascinano, come l'astronomia e i voli spaziali. Poi ci sono temi nei quali l'informazione si fa largo a fatica tra questioni etiche e paure: è accaduto vent'anni fa, quando gli italiani sono stati chiamati al referendum sull'energia nucleare, e più recentemente con il «no» dell'Italia alla ricerca sulle cellule staminali embrionali. Due «no» le cui conseguenze sono andate ben oltre le posizioni ideologiche, investendo questioni di primaria importanza per l'economia del paese così come la capacità della ricerca italiana di essere competitiva a livello internazionale.

Mediatore di questa comunicazione è stato il giornalismo scientifico, con le strategie di comunicazione e i linguaggi elaborati nel corso della sua storia, cominciata poco più di due secoli fa. Raccontare la scienza a un pubblico di non specialisti è diventato un mestiere a partire dalla fine del Settecento, e dalle primissime riviste divulgative è nata una tradizione viva ancora oggi. In questi due secoli, le riviste di divulgazione scientifica hanno conosciuto successi e fallimenti. A volte hanno avuto una vita breve e difficile; in altri casi, molto più rari, sono vissute per decenni, superando cambiamenti epocali e guerre. Hanno raccontato la scienza con linguaggi e strategie diversi, dando della conoscenza scientifica immagini molto differenti. C'è stato un modello vincente? Una strategia di comunicazione più efficace delle altre? La loro storia può aiutare a rispondere a queste domande e, forse, a capire molte delle difficoltà che la comunicazione tra scienza e società incontra oggi.

DUE SECOLI DI SCIENZA RACCONTATA. Dal 1788, quando iniziarono le pubblicazioni della «Biblioteca fisica d'Europa», in Italia sono nate 80 riviste di divulgazione scientifica, con alterne fortune. Il grande boom si registrò con la Rivoluzione Industriale, nella seconda metà dell'Ottocento, quando circolavano contemporaneamente 13 riviste.

L'ASTRONOMIA DI FLAMMARION

LA SCIENZA PER TUTTI

Divulgatore di grandissimo successo, il francese Camille Flammarion scrisse alcune «guide di astronomia «popolare» che furono pubblicate in Italia a puntate da «La scienza per tutti» nei primi anni della sua storia.

Nel corso del XVIII secolo, esempi di giornalismo scientifico nascevano in tutta l'Europa, a partire dai paesi in cui si era affermata maggiormente la scienza moderna, e in cui erano evidenti le ricadute che le scoperte scientifiche e tecnologiche avevano sulla società. Anche in Italia venivano fondate riviste di scienza che si rivolgevano a una vasta categoria di lettori di cui facevano parte imprenditori, uomini di cultura, donne e semplici curiosi. Il primo esempio risale al 1788: da allora sono stati pubblicati in Italia 80 periodici divulgativi. La maggior parte di essi ha avuto una vita molto breve, appena uno o due anni, in moltissimi casi per problemi finanziari, affrontati faticosamente da un editore che spesso era lo stesso fondatore e autore del periodico.

In questo modo avventuroso, pionieristico, a volte quasi eroico, sono nate le prime riviste divulgative italiane. All'inizio non c'erano differenze tra Nord e Sud: basti pensare che i primi due periodici di questo tipo sono stati pubblicati, a distanza di una manciata di anni, in Lombardia e nel Regno delle Due Sicilie. Nell'arco di qualche decennio la loro distribuzione geografica è diventata decisamente disomogenea, con la maggior parte delle riviste (49, pari al 60,9 per cento) nate nel Nord, 22 al Centro e appena 9 nel Sud e nelle Isole.

Altrettanto disomogenea è stata la loro distribuzione nel tempo, con periodi di attività intensa e periodi in cui queste pubblicazioni sono state rare. Un vero e proprio boom dell'editoria divulgativa, periodici compresi, c'è stato negli anni della Rivoluzione Industriale: nel ventennio 1878-1897 circolavano 13 riviste, e 11 nel decennio successivo (1898-1907).

A questo intenso periodo è seguito un sensibile rallentamento delle pubblicazioni, fino a una totale paralisi negli anni della prima guerra mondiale. Il primo segno di ripresa arrivò nel 1924, con la nascita di «Scienza e tecnica pratica»; nel 1928 seguì «Natura» e nel 1929 fu la volta di «Scienza e vita». Da allora, la pubblicazione di periodici divulgativi proseguì con un ritmo costante, anche durante la seconda guerra mondiale.

Negli anni venti per le riviste divulgative fu quasi un ripartire da zero, visto che soltanto «La Scienza per tutti» era riuscita – con non poche difficoltà – a superare indenne gli anni della guerra, e che dal 1914 al 1924 era stata l'unico periodo di divulgazione scientifica pubblicato in Italia. Il ritmo delle nuove testate accelerò all'inizio degli anni trenta, con «Mondo d'oggi» (1934) e «Sapere» (1935). Nel decennio 1948-1957 venivano pubblicate 13 testate divulgative, ma il record per il più alto numero di riviste divulgative in circolazione si è registrato nel decennio 1978-1987, con ben 17 testate in circolazione.

Contenuti e linguaggi

Questi dati bastano per capire che il panorama della divulgazione in Italia si è trasformato più e più volte in modo radicale nel giro di pochi anni. E insieme alle testate sono cambiati i contenuti. Medicina e industria, per esempio, si sono affermate fin dall'inizio come i temi più interessanti, a

IL PROGRESSO

Anno XII. 15 GENNAIO 1884 N. 1.

RIVISTA QUINDICINALE ILLUSTRATA DELLE NUOVE INVENZIONI E SCOPERTE

NOTIZIE SCIENTIFICHE, INDUSTRIALI, COMMERCIALI E VARIETÀ INTERESSANTI

SI PUBBLICA IL 15 ED IL 30 D'OGNI MESE

LO SPETTRO SOLARE

NEL SEGNO DEL PROGRESSO. Il linguaggio delle riviste nate alla fine dell'Ottocento è tutto teso alla promozione della scienza e della tecnologia come veicoli di innovazione per tutta la società.

LA SCIENZA PER TUTTI

Rivista quindicinale di volgarizzazioni scientifiche

SOMMARIO

Tutte le immagini di questo articolo provengono dalla collezione privata dell'autrice

L'AUTORE

ENRICA BATTIFOGLIA, giornalista scientifica, lavora per l'agenzia ANSA. L'articolo è una sintesi della sua tesi di laurea in filosofia del linguaggio, discussa all'Università «La Sapienza» di Roma con il linguista Tullio De Mauro e il fisico Carlo Bernardini.

cui è sempre stato dedicato lo spazio maggiore da tutte le riviste e in tutti i periodi.

Anche l'astronomia ha sempre avuto molto spazio, conservando nell'arco di due secoli il posto conquistato tra le dieci discipline più divulgate, e lo stesso è accaduto per la fisica. Per gli altri, nei diversi periodi storici è avvenuto un ricambio costante: nuove discipline si sono sostituite ad ambiti ormai superati e non più legati all'attualità né dal punto di vista scientifico né da quello economico, sociale e culturale. Nuove aree tematiche sono emerse e si sono affermate, come l'ambiente, lo spazio e l'informatica, mentre altre sono scivolate in secondo piano, come è accaduto a partire dal secondo dopoguerra per argomenti come l'elettricità, il telefono e il telegrafo, che si erano guadagnati uno spazio notevole nel XIX secolo.

Cambiamenti analoghi sono avvenuti nel linguaggio del giornalismo scientifico. Lo stile più vicino alla cronaca è stato il primo ad affermarsi, e si è proposto come un modello efficace e sicuramente vincente, basato sull'informazione e l'attualità. Anche la dimensione spettacolare della scienza è stata proposta con successo, specialmente se accompagnata da immagini. Vita più difficile hanno invece avuto le riviste che si proponevano ai lettori con uno scopo dichiaratamente educativo, puntando alla formazione più che all'informazione.

Stili diversi sono stati usati nel tempo anche per veicolare differenti immagini della scienza: da strumento di progresso a portatrice di luce, da religione dell'avvenire a fonte di conoscenze straordinarie e meravigliose.

Piccole industrie, grandi scienziati

In Italia, come nel resto d'Europa, il giornalismo scientifico è nato nel periodo in cui la cultura illuministica trasformava la scienza in un fenomeno di moda, al punto da farla entrare nei salotti esaltandone il carattere spettacolare. Sono celebri le dimostrazioni pubbliche dell'abate Nollet, in cui la bottiglia di Leida diventava l'occasione per stupire e divertire con il nuovo fenomeno dell'elettricità. Perfino le leggi dell'ottica formulate da Newton uscivano dalle Accademie per intrattenere e incuriosire un pubblico di persone colte, ma non legate alle università. A trasformare con disinvoltura la scienza in un ingrediente irrinunciabile delle serate mondane erano uomini come Francesco

RECORD DI LONGEVITÀ. Tra le riviste nate con la Rivoluzione Industriale, «La Scienza per tutti» è stata la più duratura, uscendo quasi ininterrottamente per 66 anni, dal 1877 al 1943. La rivista più longeva in assoluto, però, è «Sapere», fondata come quindicinale nel 1935 e pubblicata ancora oggi, con periodicità bimestrale, con la direzione di Carlo Bernardini.

Algarotti, il cui libro *Newtonianismo per le dame* (1737) ebbe un ruolo di primo piano nella diffusione delle teorie newtoniane in Italia e nel resto d'Europa, con ben 31 edizioni e traduzioni in sei lingue: tanto da diventare rapidamente «uno dei libri più letti di un autore italiano del XVIII secolo» e da entrare a far parte dell'*Indice dei libri proibiti* di Benedetto XIV.

I libri divulgativi di Algarotti non erano un fenomeno isolato: molti suoi contemporanei hanno avuto un ruolo importante in Italia nel diffondere le nuove idee scientifiche, come il medico veneziano Eusebio Sguario, autore di *La chimica per le donne*, Giuseppe Compagnoni e Antonio Cagnoli, autore di celebri almanacchi.

Negli ultimi decenni del Settecento, inoltre, le notizie scientifiche erano diventate interessanti anche per le applicazioni sempre più numerose e importanti che avevano nella medicina, nell'agricoltura e nell'industria. I primi a riservare uno spazio del tutto particolare a queste «informazioni utili» furono il medico, chimico e naturalista Luigi Valentino Brugnatelli (1761-1818) e un giovane medico e industriale abruzzese, Vincenzo Comi (1764-1839). Il primo lo fece pubblicando nel 1788 la «Biblioteca fisica d'Europa», primo esempio di periodico divulgativo, teso a «far conoscere al pubblico italiano le principali memorie degli scienziati d'oltralpe», il secondo fondando nel 1792 il «Commercio scientifico d'Europa con il Regno delle Due Sicilie».

Tutte e due queste prime riviste utilizzavano come fonti principali le memorie dei maggiori scienziati europei, che trasformavano in notizie, selezionate e scritte in modo da informare sugli aspetti più attuali e utili del sapere scientifico e delle novità tecnologiche. Brugnatelli e Comi si rivolgevano agli uomini di scienza e agli industriali per avere informazioni e aggiornamenti su argomenti ritenuti di pubblica utilità.

Nascevano così le prime cronache scientifiche. Il loro obiettivo era fornire notizie che permettessero ai non scienziati di essere aggiornati sulle novità teoriche e pratiche sulla scienza e la tecnologia, la medicina, l'igiene e l'industria, immediatamente seguite da notizie di fisica e storia naturale. In ogni caso, a prevalere erano la cronaca e l'attualità, sia che si dessero aggiornamenti su tecniche chirurgiche e nozioni di igiene, sia che si descrivessero, con tanto di illustrazioni, i tentativi di volo fatti con un paio di ali di tela, nel bel mezzo del parco di Vienna, da un avventuroso inventore.

Pur avendo aperto la strada, le prime due riviste italiane rimasero a lungo dei casi isolati. Brugnatelli fondò in seguito un altro periodico, il «Giornale di Fisica», ma per molti anni la pubblicazione di nuove riviste divulgative restò un fenomeno sporadico. Anche perché in questo periodo iniziale, che va dal 1788 alla prima metà dell'Ottocento, i lettori dei periodici divulgativi erano ancora un gruppo molto ristretto di uomini di cultura.

La scienza che vale

Nella seconda metà del secolo XIX le cose cambiarono notevolmente, e la pubblicazione di riviste scientifiche divulgative, da fenomeno sporadico, registrò un vero e proprio boom, un'età dell'oro che durò fino alla vigilia della prima guerra mondiale.

Sulla scia di un'intensa attività scientifica segnata dall'affermarsi di nuove teorie, come l'evoluzionismo, e dell'avanzare della tecnologia, negli anni della Rivoluzione Industriale in Italia i periodici di scienza popolare venivano pubblicati a ritmo vertiginoso, tanto che in 44 anni comparvero ben 28 nuove riviste, pari al 35,4 per cento del totale delle riviste di questa categoria pubblicate in oltre due secoli. Negli stessi anni, negli Stati Uniti si affermava «Scientific American», nata nel 1845 e destinata a diventare la rivista divulgativa più longeva del mondo, e in Francia fioriva un'editoria divulgativa destinata a diventare il modello di riferimento di molti periodici italiani di scienza popolare, con «La Science en Famille» o «La Nature».

Negli anni dell'affermarsi dell'industria e delle grandi esposizioni, la tecnologia saliva in primo piano nella maggior parte delle riviste divulgative. Uno spazio importante era assegnato anche a notizie sulla fisica e sull'energia, mentre l'astronomia continuava a esercitare un grande fascino, insieme alla zoologia, ai viaggi e alla natura. Contemporaneamente, l'attenzione delle riviste divulgative dell'Ottocento si rivolgeva a fenomeni di attualità e dal grande impatto sociale, come la nascita dell'industria e l'arrivo delle macchine (dai grandi impianti industriali alle macchine per cucire o per maglieria), insieme all'elettricità, al telegrafo e al telefono. Periodici come «Il Progresso» e la «Rivista Scientifico-Industriale» presentavano quelle innovazioni come veri simboli del progresso che coinvolgevano l'intera società, avevano ripercussioni dirette sulla vita quotidiana ed entravano nelle case rivoluzionando abitudini e tradizioni.

Che si parlasse di industria, di fisica, di elettricità o di chimica, a essere privilegiato era soprattutto il valore applicativo delle scoperte e delle innovazioni. Si aggiornavano i lettori sulle novità, e nello stesso tempo si fornivano loro le informazioni indispensabili per comprenderle, spiegando per esempio i principi del funzionamento di elettricità, telegrafo, telefono, mezzi di trasporto. L'attenzione per l'informazione tecnica era così forte da portare alla pubblicazione di riviste interamente dedicate ai prodotti e alle applicazioni della ricerca.

Divulgatori o missionari?

La cronaca e l'informazione non erano gli unici modelli di comunicazione della scienza in circolazione nella seconda metà del XIX secolo. Alcune testate, come «La Scienza e la fede», cercarono di avvicinare i lettori adottando la formula del dialogo tra insegnante e studente, nel quale il primo rispondeva con estrema semplicità ai dubbi e alle domande sollevati dal secondo. Si trattava tuttavia di un espediente al quale si ricorreva soprattutto nell'analizzare argomenti filosofici, come il rapporto tra religione e conoscenza scientifica. La modalità del dialogo è stata una delle meno utilizzate nella storia dei periodici divulgativi, benché si sia più volte riaffacciata, anche dopo lunghi intervalli di tempo. Per esempio, a distanza di quasi un secolo dalla pubblicazione di «La Scienza e la fede», fra gli anni quaranta e cinquanta del Novecento, la rivista «Scienza e lavoro» continuava a prendere per mano il lettore servendosi del dialogo per introdurre e spiegare nuovi termini e soprattutto nuovi concetti.

Sempre nella seconda metà dell'Ottocento alcune riviste divulgative preferivano introdurre il pubblico ai temi scientifici attraverso brevi racconti, spesso ispirati alla storia degli scienziati e delle scoperte, il cui stile era molto simile a quello di un romanzo. Molte volte quegli articoli erano ripresi dai libri di successo scritti da grandi divulgatori francesi, come *La vita normale e la salute*, di Emile Rengade, o *La terra e il cielo*, di Camille Flammarion, le cui opere furono pubblicate a puntate da «La Scienza per tutti» nei primi anni della sua storia. La stessa rivista pubblicò la serie di episodi dedicati ai «Martiri della scienza», in cui aneddoti della vita di celebri scienziati erano raccontati con lo stile di un romanzo d'appendice. La scienza aveva i suoi martiri perché, secondo alcuni divulgatori, era in fondo una sorta di religione:

«la religione dell'avvenire», come la definì in un editoriale del 1884 il direttore della rivista «La Natura», Paolo Mantegazza. Era una divulgazione popolare dalla forte caratterizzazione ideologica, spesso orientata alla difesa delle nuove teorie del positivismo.

Più che di invenzioni e scoperte, nelle riviste che seguivano questo orientamento si parlava della scienza come di un'arma per lottare contro pregiudizi e superstizioni, capace di trovare spiegazioni razionali e fondate su dati reali anche di fenomeni apparentemente inspiegabili. La metafora della scienza come di una guerra capace di servirsi di strumenti benefici per scopi nobili era diffusa soprattutto negli ultimi decenni del XIX secolo. «Slanci eroici di geni» erano i contributi degli scienziati, visti come «eroi del pensiero» e «militi della scienza» al servizio di una «sublime crociata», tanto da sfidare ogni pericolo e andare «sereni incontro al martirio».

La metafora della scienza come strumento di conquista ebbe particolare successo in molti periodici divulgativi del ventennio fascista, che vedevano nella ricerca un'attività fondamentale alla vita del paese e un «potente» sostegno per l'industria. «Mai come ora la ricerca scientifica è divenuta necessaria per l'economia, per la vita sociale, per la difesa del Paese», osservava Guglielmo Marconi nel *Messaggio agli scienziati d'Italia* pronunciato in occasione della riunione della Società italiana per il progresso delle scienze (Sips) che si tenne a Napoli nell'ottobre 1934.

Il trionfo della certezza

L'immagine di una scienza forte e portatrice di certezze, oltre che di ricchezza, era emersa per la prima volta in modo evidente in alcune riviste divulgative pubblicate negli anni della Rivoluzione Industriale. Lo strumento per veicolare l'immagine dominante della scienza era l'uso di un linguaggio pieno di affermazioni, basato sull'uso del tempo presente e del futuro. Tuttavia, l'aspetto più evidente di questa tendenza si affermò nel ventennio fascista, con il ricorso molto frequente all'uso dell'affermazione nelle riviste scientifiche. Il linguaggio tipico della propaganda portò anche all'uso di slogan in riviste che, come «Scienza e tecnica» (1937), avevano un ruolo ufficiale

in quanto rappresentative di associazioni scientifiche nazionali.

L'uso dell'affermazione continuò comunque a esercitare un certo fascino su una parte della stampa divulgativa del secondo dopoguerra. La paura di una scienza al servizio di armi micidiali, come la bomba atomica, fu più volte esorcizzata con le immagini positive di un futuro in cui la tecnologia era l'ossatura di una società progredita. «Nel 2000 saremo giovani a 100 anni», assicurava «La Scienza Illustrata» nel 1955: uno dei tanti sogni di allora sul mitico 2000, nel quale si immaginavano anche locomotive a propulsione nucleare e automobili in grado di volare e navigare («Scienza e vita», 1965).

La scienza come attività meravigliosa e capace di suscitare stupore era comparsa per la prima volta verso gli anni ottanta del XIX secolo. «Stupefacenti», «incredibili» e «prodigiosi» erano gli aggettivi utilizzati da «Scienza e vita» nel comunicare i risultati della ricerca, e ancora nel 1937 il senso di stupore nei confronti della scienza tornava a manifestarsi pienamente in un vero e proprio rotocalco divulgativo, «Il Giornale delle Meraviglie», ricchissimo di illustrazioni.

La macchia dell'atomica e la Big Science

Come era accaduto dopo la Grande Guerra, anche nel secondo dopoguerra l'attività divulgativa registrò una ripresa. Nel 1946, per esempio, nacquero

quattro nuove testate («Scienza popolare», «Natura e vita», «Historia naturalis», «Scienza e lavoro»), e nel 1949 addirittura cinque («Selezione scientifica», «Scienza e meccanica popolare», «Scienza e vita», «La Scienza Illustrata», «L'Illustrazione scientifica»). L'elemento comune a tutte era il dibattito sul valore della conoscenza scientifica e sui possibili usi che di essa possono fare scienziati o politici, scatenato dalla percezione della bomba atomica come una macchia indelebile nella storia della scienza e dell'umanità.

Dopo lo shock dell'atomica, almeno tre riviste lanciarono un appello in difesa di una scienza in grado di garantire la fratellanza universale. Più che dare informazioni, l'obiettivo della stampa periodica divulgativa in questo periodo era riscattare il sapere scientifico, offrendone un'immagine positiva, al servizio della pace. La seconda guerra mondiale segnò anche l'esordio della Big Science, ossia di progetti di ricerca così complessi e ambiziosi da richiedere investimenti molto consistenti e da rendere indispensabile la collaborazione internazionale. Imprese di questa mole portarono la necessità di una maggiore circolazione delle notizie scientifiche. La ricerca abbandonava la fase artigianale per entrare in quella dei progetti su larga scala, impossibili senza la collaborazione dei più grandi laboratori del mondo. Si era aperta la via a imprese come lo sbarco sulla Luna, la costruzione dei grandi acceleratori di particelle e dei computer, o ancora il Progetto genoma umano che ha sequenziato il codice genetico dell'uomo.

Come nella seconda metà dell'Ottocento, nel secondo dopoguerra si avvertiva la portata innovativa della scienza e della tecnologia nella vita di tutti i giorni. Nella società civile cominciarono a entrare innovazioni tecnologiche di origine militare, come il radar e il DDT; nel 1946 entrò in funzione negli Stati Uniti il primo calcolatore, e il transistor aprì una nuova era dell'elettronica. Nel 1953 James Watson e Francis Crick descrivevano la struttura del DNA e nel 1957 lo Sputnik inaugurava la corsa allo spazio. A questi grandi cambiamenti in Italia non corrispose però un forte aumento di interesse per la divulgazione scientifica, che rimase costante, con quattro nuove testate nate negli anni cinquanta e tre negli anni sessanta, tra cui anche «Le Scienze».

PER APPROFONDIRE

GOVONI P., *Un pubblico per la scienza*, Carocci, Roma, 2002.

BARBAGLI F. e BRUGNATELLI L.V., in *Poeti, scienziati, cittadini nell'ateneo pavese tra riforme e rivoluzione*, Pavia, 2000.

RUPERT HALL A., *La matematica, Newton e la letteratura*, in R. Cremante e W. Tega (*a cura*), *Scienza e letteratura nella cultura italiana del Settecento*, Il Mulino, Bologna, 1984.

CASINI P., *Les débuts du newtonianisme en Italie*, in «Dixhuitième Siècle», 1978.

ALGAROTTI F., *Newtonianismo per le Dame, ovvero Dialoghi sopra la luce i colori*, in «Opere», Cremona, 1778.

UN NUOVO SGUARDO SULLA SCIENZA. Con il secondo dopoguerra nascono molte nuove riviste, tra cui «La Scienza Illustrata» e «Scienza e vita» (entrambe del 1949). Nel settembre 1968 sarà la volta di «Le Scienze», di cui riportiamo qui a fianco la prima copertina.

Una ripresa significativa della divulgazione scientifica ci fu alla fine degli anni settanta, quando le riviste di scienza popolare beneficiarono di cospicui finanziamenti e contemporaneamente nell'opinione pubblica si diffuse una nuova attenzione alla scienza. Per questi motivi, gli anni ottanta sono stati il decennio più ricco di nuove riviste di tutto il secolo, con la nascita di ben 11 testate dal 1980 al 1988. Un momento d'oro proseguito nei primissimi anni novanta, e che, dopo una fase di stallo, ha visto una leggera ripresa a partire dal 1997.

Tuttavia, la pubblicazione di nuove riviste divulgative è rimasta un evento relativamente raro, e i periodici scientifici hanno spesso navigato in cattive acque, tanto che molti sono stati costretti a chiudere. L'attenzione per l'informazione scientifica non è aumentata in modo apprezzabile, a eccezione di qualche leggero segnale di ripresa alla fine degli anni novanta che ancora oggi lascia sperare in un nuovo interesse del pubblico per la scienza. Tuttavia nel nostro paese c'è ancora molto da fare perché la ricerca scientifica venga percepita non soltanto come un argomento di interesse culturale, ma come un fattore essenziale per lo sviluppo economico e sociale.

La lezione delle più longeve

Delle 80 riviste divulgative pubblicate in Italia nell'arco di due secoli, le più longeve sono state «La Scienza per tutti», uscita quasi ininterrottamente per 66 anni, dal 1877 al 1943, la «Rivista Scientifico-Industriale», vissuta quarant'anni, dal 1869 al 1909, *Il Progresso* (29 anni, dal 1873 al 1902) e due testate tuttora in vita: «Sapere», nata nel 1935, e «Le Scienze», nata nel 1968. Complessivamente, solo il 33 per cento delle riviste di questo tipo è riuscito a vivere per oltre dieci anni.

Oltre alla durata, queste testate hanno in comune una forte aderenza all'attualità e alla cronaca. «La Scienza per tutti», per esempio, dopo un periodo iniziale nel quale si limitò a riproporre brani dei libri dei grandi divulgatori francesi, si indirizzò decisamente verso articoli originali sui progressi dell'industria e le novità tecnologiche e scientifiche. «Sapere» ai suoi inizi era un quindicinale capace di affrontare la scienza dal punto di vista della cronaca, e di dare spazio, nello stesso tempo, a interventi diretti di grandi scienziati, come Enrico Fermi. E anche «Le Scienze» ha saputo cambiare nel tempo, privilegiando l'attenzione al nuovo e ai temi scientifici più aggiornati e vicini alla cronaca.

È nell'attenzione all'attualità, nella scelta di temi vicini alla cronaca e nella capacità di adattare il linguaggio a quello del pubblico la spiegazione della longevità di questi periodici? Pur considerando i problemi reali relativi al rapporto con gli editori e alle disponibilità finanziarie, è molto probabile che la concretezza dello stile e dei temi, insieme all'aderenza alla cronaca e all'informazione aggiornata siano fattori cruciali nel determinare il successo delle riviste di scienza popolare.

Controllare il dolore

I progressi compiuti nella comprensione dei meccanismi cellulari e molecolari che trasmettono i segnali di dolore indicano nuovi bersagli per farmaci analgesici che potrebbero alleviare diverse forme di dolore, tra cui anche quelle poco controllabili dalle terapie di oggi

di Allan I. Basbaum e David Julius

Pulsante, insistente, lancinante, martellante, penetrante. Il dolore può presentarsi in molte forme, tutte sgradevoli. Ma che si tratti del bruciore provocato dal passaggio di un calcolo renale o dalla morsa opprimente causata da un'emicrania, tutti i tipi di dolore hanno una caratteristica in comune: chi ne soffre vorrebbe eliminarlo. Eppure gli analgesici ancora oggi più diffusi sono sostanzialmente rimedi popolari che hanno funzionato per secoli: la morfina e gli altri oppiacei derivano dal papavero da oppio, l'aspirina si ricava dalla corteccia del salice. Per quanto siano in grado di alleviare il dolore, ognuno di questi farmaci ha dei limiti. L'aspirina e gli altri antinfiammatori non steroidei (FANS) non riescono ad alleviare le forme di dolore più gravi. Persino gli oppiacei, che in genere sono i farmaci più aggressivi, non sono efficaci su tutti. Senza contare che possono causare gravi effetti collaterali e che, siccome i pazienti tendono a divenire tolleranti nei loro confronti, è necessario aumentare il dosaggio per ottenere un certo sollievo. Negli ultimi vent'anni i neurobiologi hanno acquisito moltissime conoscenze sui circuiti cellulari e sulle molecole specializzate che veicolano i segnali dolorosi. Oggi queste conoscenze vengono sfruttate per mettere a punto nuove strategie in grado di gestire meglio il dolore, causando meno effetti collaterali. In effetti, gli approcci attualmente allo studio sono assai numerosi, e non c'è modo di prenderli tutti in esame in questa sede.

Pasieka/SPL/Grazia Neri

QUEL BRIVIDO LUNGO LA SCHIENA. Sono i neuroni del midollo spinale, che ricevono i segnali dalle cellule specializzate nel rispondere agli stimoli dolorosi, a trasportare i messaggi fino al cervello.

Particelle di fuoco

Nel XVII secolo Cartesio elaborò una teoria per spiegare la percezione del dolore. Credeva che un pizzicotto, un colpo, una puntura attivassero un circuito nervoso che poi faceva suonare un campanello di allarme nel cervello. Immaginate di scottarvi un piede. Cartesio credeva che «particelle di fuoco in rapido movimento» creassero una perturbazione che «si propagava lungo il nervo fino a raggiungere il cervello».

Non era lontano dal vero. Di solito il dolore si forma alla periferia: sulla pelle, in un organo o in qualsiasi altra sede esterna al sistema nervoso centrale, cioè cervello e midollo spinale. Quando urtiamo qualcosa con un dito o ci appoggiamo al fornello bollente attiviamo cellule nervose, i nocicettori, che rispondono a stimoli che danneggiano i tessuti, come temperature estreme, pressione meccanica o sostanze chimiche prodotte in risposta a una ferita o a un'infiammazione.

I nocicettori hanno due rami: uno, deputato a percepire la sensazione, che si proietta verso la periferia, dove innerva piccole porzioni di tessuto, e uno che si protende nel midollo spinale. Il corpo cellulare del neurone, che risiede in una struttura esterna al midollo spinale, si trova a metà strada fra i due. Quando le molecole che rilevano i segnali nel ramo periferico percepiscono un agente nocivo fanno partire un impulso che percorre il ramo centrale fino a raggiungere un'area del midollo spinale chiamata corno dorsale. Qui i nocicettori liberano neurotrasmettitori che attivano i neuroni del corno dorsale inducendoli a trasmettere il messaggio al cervello. Spesso sono chiamati «neuroni che rilevano il dolore», ma i nocicettori si limitano a segnalare la presenza di stimoli potenzialmente pericolosi. È il cervello a interpretare il segnale come doloroso.

Non tutti i tipi di dolore suscitano preoccupazione. Il dolore acuto che accompagna una lussazione o un'abrasione, per esempio, ha una funzione protettiva: spinge l'organismo a evitare ulteriori danni. Questo genere di dolore tende a essere transitorio e ad attenuarsi col tempo. Il dolore cronico che affligge molti pazienti, invece, non riesce a scomparire ed è difficile da trattare. In molti casi, il problema si presenta perché la ferita o l'infiammazione che scatena il malessere non si estingue. I dolori causati dall'artrite derivano da un'infiammazione in atto, mentre il tormento che accompagna un tumore invasivo deriva, in gran parte, dalle lesioni a carico dei tessuti e dall'infiammazione.

In altri casi il dolore persistente è neuropatico, cioè deriva da un danno che interessa le cellule nervose. Può svilupparsi quando i neuroni del sistema nervoso centrale sostengono il danno causato da sclerosi multipla, da un infarto o da una lesione al midollo spinale, per esempio. In alternativa, può derivare da lesioni a carico di neuroni periferici. Soffrono di dolore neuropatico tutti coloro che hanno subito un'amputazione e percepiscono un dolore persistente in un arto che ormai non c'è più (la cosiddetta sindrome dell'arto fantasma), e coloro che, per anni, continuano ad avvertire una sensazione di bruciore cutaneo dopo che un'infezione da herpes virus si è attenuata (la nevralgia posterpetica). Il persistere di questo tipo di dolore non è sintomatico di ferite in corso o di un'altra malattia: è un disturbo del sistema nervoso, e deve essere seguito da uno specialista del dolore.

Dolore senza fine

Uno dei principali fattori comuni alle persone che soffrono di dolore intrattabile è l'ipersensibilità agli stimoli, che può presentarsi sotto forma di iperalgesia (una reazione eccessiva a stimoli normalmente dolorosi), o di allodinia (dolore dovuto a uno stimolo normalmente innocuo). Nelle persone che soffrono di allodinia, persino la lieve pressione dei vestiti contro la pelle o un'articolazione che si flette possono causare una sofferenza insopportabile.

Questa ipersensibilità deriva da modifiche molecolari o strutturali che si verificano nelle cellule nervose. In periferia, per esempio, molecole pro-infiammatorie possono rendere i nocicettori eccessivamente reattivi ai segnali di dolore. Le molecole infiammatorie possono persino indurre i nocicettori a iniziare a produrre segnali anche in assenza di uno stimolo.

Anche i cambiamenti del sistema nervoso centrale che determinano iperattività dei circuiti di trasmissione del dolore possono indurre sensibilizzazione. Uno di questi cambiamenti, potenzialmente duraturi, comprende l'esposizione di un maggior numero di recettori che rispondono ai neurotrasmettitori rilasciati dai nocicettori, e potrebbe persino provocare un riarrangiamento delle connessioni nervose e la perdita di cellule nervose che inibiscono gli stimoli dolorosi.

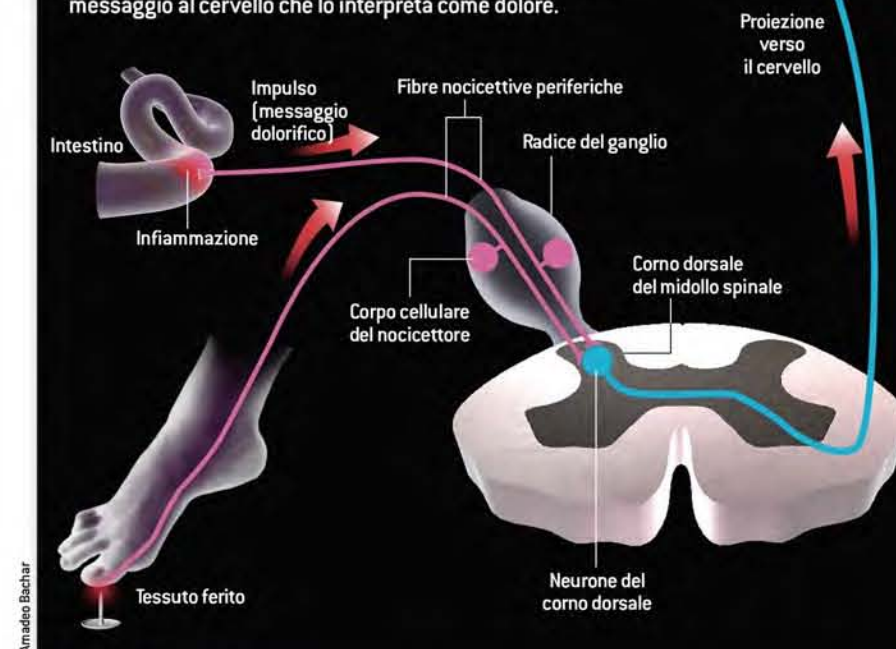
Indipendentemente da quali sono i processi malfunzionanti, il dolore in atto può determinare la sensibilizzazione, e dunque esacerbare e prolungare la sofferenza. Pertanto l'obiettivo di molti scienziati alla ricerca di nuovi analgesici è migliorare l'iperalgesia e l'allodinia. Nel frattempo, i pazienti devono capire che il dolore non deve essere sopportato stoicamente: ci vuole un trattamento aggressivo, per prevenire un'ulteriore sensibilizzazione.

I pazienti devono capire che il dolore non deve essere sopportato stoicamente: ci vuole un trattamento aggressivo

In sintesi/*Alleviare il dolore*

- Cellule nervose specializzate, i nocicettori, rispondono a stimoli nocivi trasmettendo un messaggio ai neuroni del midollo spinale: questi, poi, trasportano il segnale fino al cervello.
- I nocicettori e altri neuroni del circuito del dolore hanno speciali molecole in grado di individuare gli stimoli che procurano dolore. Queste molecole potrebbero essere il bersaglio per sviluppare medicinali in grado di alleviare il dolore, causando meno effetti collaterali di quelli provocati dai farmaci attualmente in commercio.

COME PERCEPIAMO IL DOLORE

Il circuito del dolore, raffigurato qui in forma semplificata, si estende dalla periferia dell'organismo – dalla pelle e dagli altri tessuti esterni al sistema nervoso centrale – fino al midollo spinale e al cervello. Stimoli dolorosi attivano particolari cellule nervose che percepiscono il dolore, i nocicettori (*in viola*), i quali inviano gli impulsi lungo gli assoni, trasportando il messaggio *di allerta* alle cellule nervose nella regione del corno dorsale del midollo spinale (*in blu*). Queste cellule, a loro volta, trasmettono il messaggio al cervello che lo interpreta come dolore.

Cominciare dall'inizio

La ricerca di nuovi analgesici si è orientata allo studio del luogo da cui hanno origine i segnali: la periferia. Alcune molecole usate dai nocicettori per percepire stimoli dannosi sono rare in altre parti del corpo. Bloccandole, si dovrebbero fermare i segnali dolorifici senza perturbare altri processi fisiologici e, quindi, senza provocare effetti collaterali indesiderati.

La maggior parte dei rimedi più diffusi fa sentire i propri effetti benefici soprattutto alla periferia. Quando un tessuto è ferito, molte cellule presenti in loco secernono prostaglandine, che agiscono sui rami dei nocicettori deputati alla percezione del dolore abbassandone la soglia di attivazione. L'aspirina e i FANS inibiscono l'attività di una famiglia di enzimi (le cicloossigenasi) che le cellule usano per generare le prostaglandine che causano il dolore. Questi farmaci danno sollievo alle più comuni forme di sofferenza, ma inibiscono la produzione di prostaglandine in altre parti del corpo, provocando effetti collaterali come bruciore di stomaco, diarrea e ulcere che possono impedirne l'uso prolungato e limitare il dosaggio.

Per ridurre le conseguenze a livello gastrointestinale, le case farmaceutiche hanno messo a punto una famiglia di farmaci il cui bersaglio è costituito dalla cicloossigenasi-2 (COX-2). Siccome, di norma, la COX-2 non è attiva nello stomaco e nel tratto intestinale, il blocco della sua attività non dovrebbe causare i disturbi provocati dai FANS tradizionali. Resta però da appurare se abbiano davvero un effetto più blando sullo stomaco. Alcuni farmaci hanno però posto problemi. Come il Vioxx, prescritto per alleviare il dolore causato dall'artrite, che è stato ritirato dal mercato quando si è scoperto che aumentava il rischio di infarto al miocardio e ictus. E altri inibitori della COX-2 sono sotto esame per individuare effetti indesiderati.

La scoperta di bersagli presenti quasi esclusivamente sui nocicettori offre un'opportunità per sviluppare farmaci che agiscono selettivamente alleviando il dolore.

Ecco a voi la salsa

Un bersaglio particolarmente allettante è il recettore per la capsaicina, un canale ionico presente sulla membrana di molti nocicettori che risponde alla capsaicina, la molecola piccante del peperoncino, ma anche al dolore causato dal calore, e ai protoni (gli ioni idrogeno, che conferiscono acidità alle sostanze). I protoni, tra l'altro, sono insolitamente abbondanti nei tessuti infiammati. In presenza di queste sostanze o di temperature sopra i 43 gradi, il canale ionico permette a ioni sodio e calcio di entrare nel nocicettore, stimolandolo a produrre un segnale che si traduce nella sensazione di bruciore indotta da calore, infiammazione e cibi piccanti.

Pertanto le sostanze che inibiscono i recettori per la capsaicina dovrebbero attenuare il dolore infiammatorio. In esperimenti su animali da laboratorio, questi antagonisti della capsaicina sono riusciti ad alleviare il dolore provocato dall'ambiente acido che circonda i tumori che hanno prodotto metastasi e danneggiato il tessuto osseo. Molte case farmaceutiche stanno cercando di mettere a punto antagonisti del recettore per la capsaicina.

Ma non è finita qui. In alcuni casi i recettori per la capsaicina, opportunamente stimolati, possono alleviare il dolore. Creme a base di capsaicina sono prescritte per alleviare il prurito che accompagna la guarigione di una ferita post-operatoria o il danno neurale derivante da infezioni da HIV, accessi di herpes e diabete. L'esatta modalità di azione di queste pomate non è chiara, sebbene dosi modeste, col passare del tempo, possono alla fine rendere il recettore meno reattivo nei confronti di stimoli abituali, o portare all'esaurimento dei neurotrasmettitori secreti dai nocicettori.

Bloccare altri canali

Un altro tipo di molecola rinvenuta sulle terminazioni periferiche dei nocicettori sta suscitando interesse come potenziale bersaglio farmacologico. Tutti i neuroni hanno diversi tipi di canali del sodio. L'apertura di questi canali, che si verifica in risposta ai cambiamenti di voltaggio tra l'interno e l'esterno della membrana neuronale, genera gli impulsi che trasmettono i messaggi da un neurone al successivo. Oggi, grazie agli anestetici locali che inat-

tivano temporaneamente i canali è possibile trattare molte forme di dolore. Il problema è che questi anestetici devono essere applicati nel punto sofferente: l'inattivazione dei canali in tutto il sistema nervoso potrebbe essere fatale.

I neuroni che percepiscono il dolore sono dotati di una sottoclasse di canali del sodio, i canali TTX-resistenti, assenti nel sistema nervoso centrale. Perciò i ricercatori sperano di poter somministrare in tutto l'organismo farmaci in grado di bloccare questi canali senza effetti avversi. Agenti di questo tipo potrebbero attenuare anche la sconveniente iperattività causata da nervi periferici lesionati, alleviando alcune forme di dolore neuropatico. Finora, purtroppo, l'industria farmaceutica non è riuscita a sviluppare efficaci inibitori diretti selettivamente contro questi canali.

Un'altra possibilità è rimuovere selettivamente i canali con l'interferenza dell'RNA. Questa tecnica si basa sull'introduzione nell'organismo di minuscole molecole chiamate piccoli RNA interferenti, o siRNA. I siRNA impediscono la produzione di una proteina indesiderata inducendo la degradazione delle molecole che ne dirigono la sintesi (le molecole di RNA messaggero). I ricercatori stanno studiando l'applicabilità di questa tecnica all'uomo per il trattamento di certe patologie retiniche, ma la sfida del futuro è riuscire a trasformare l'interferenza dell'RNA in un trattamento farmacologico analgesico. Come accade anche per la terapia genica, è verosimile che per trasportare a destinazione le molecole di siRNA serva un virus, e questa esigenza ha sollevato preoccupazioni relative alla sua sicurezza. Il tempo ci dirà se questo approccio sarà praticabile.

Supponiamo che le case farmaceutiche mettano davvero a punto un analgesico che elimini in modo specifico ed efficace l'attività di una delle molecole, presenti sui nocicettori, che trasmettono il dolore. Ma un intervento simile darebbe sollievo nei casi di dolore cronico? Sbarrare un'unica via d'accesso al percorso dei segnali potrebbe non bastare.

Immaginate che si riesca a eliminare il recettore per la bradichinina, un peptide prodotto nelle infiammazioni periferiche. La bradichinina è un potente stimolante dei nocicettori, e un antagonista in grado di bloccarne i recettori potrebbe impedire l'attivazione dei nocicettori. Ma non impedirebbe ai neuroni di riconoscere altre molecole prodotte da una ferita o da un'infiammazione, come protoni e prostaglandine. Allo stesso modo, mettere fuori gioco solo i recettori per la capsaicina potrebbe non ridurre tutto il dolore mediato dai protoni, dato che in certe circostanze i protoni attivano una popolazione separata di recettori, gli ASIC (dall'inglese *Acid-Sensing Ion Channel*), sui nocicettori.

GLI AUTORI

ALLAN I. BASBAUM e DAVID JULIUS studiano i meccanismi cellulari e molecolari che sottendono al dolore. Basbaum, che ha conseguito il PhD in neuroscienze all'Università della Pennsylvania, è professore e preside del Dipartimento di anatomia all'Università della California a San Francisco (UCSF). Julius, che ha conseguito il PhD in biochimica all'Università della California a Berkeley, insegna farmacologia cellulare e molecolare alla UCSF.

I ricercatori dovrebbero mettere a punto terapie cognitive migliori per modificare la percezione del dolore

Concentrarsi sul midollo

Una strategia per aggirare il problema costituito da questa ridondanza potrebbe essere somministrare un cocktail di molecole ad attività inibitoria in grado di colpire più meccanismi dolorifici. Oppure si potrebbero scegliere come bersaglio molecole che agiscono a livello più centrale, impedendo a tutti i nocicettori – indipendentemente da qual è stato lo stimolo che li ha attivati – di trasmettere i loro segnali dolorifici ai neuroni del midollo spinale.

La morfina e gli altri oppiacei, che si legano ai recettori per gli oppiacei sulle terminazioni dei nocicettori che raggiungono il midollo spinale, sfruttano proprio questa tattica. Quando attivano i recettori, gli oppiacei impediscono il rilascio del neurotrasmettitore, bloccando la trasmissione del segnale ai neuroni del midollo spinale. Inoltre riducono la capacità dei neuroni del corno dorsale di rispondere a stimoli dolorifici. Dato che questi farmaci agiscono nel midollo spinale, in teoria dovrebbero trattare tutti i tipi di dolore. In realtà funzionano soprattutto sulle forme collegate a processi infiammatori.

Purtroppo i recettori per gli oppiacei sono presenti sui neuroni di tutto l'organismo, compresi il cervello e il sistema gastrointestinale: questo spiega perché la morfina e le altre sostanze della stessa famiglia possono produrre un'ampia gamma di effetti collaterali indesiderati. Questi problemi pongono un limite alla dose di farmaco che il paziente può assumere in sicurezza, o che un medico può prescrivere. Per di più, molti medici sono restii a prescrivere oppiacei, temendo che i pazienti sviluppino dipendenza.

La dipendenza non è comune tra i pazienti che assumono oppiacei come analgesici. Anche per evitare alcuni degli effetti collaterali, questi farmaci sono spesso somministrati direttamente nello spazio pieno di liquido che circonda il midollo spinale. E possono anche essere somministrati con un'iniezione (nel caso del dolore postoperatorio) o mediante una pompa infusionale (per il dolore cronico).

E poi ci sono le alternative agli oppiacei. Farmaci che interferiscono con i canali del calcio possono impedire il rilascio dei neurotrasmettitori dalle terminazioni dei nocicettori nel midollo spinale. Si pensa che la gabapentina, un anticonvulsivo, allevi alcune forme di dolore interagendo con una subunità specifica di certi canali del calcio. Ziconotide, un farmaco che si ricava dal veleno di un gasteropode che vive nell'Oceano Pacifico, inibisce un altro tipo di canali del calcio: i canali di tipo N.

Come accade con i recettori per gli oppiacei, anche i canali del calcio di tipo N sono diffusi in tutto il sistema nervoso. Se si somministrasse ziconotide per via sistemica, la pressione sanguigna scenderebbe bruscamente: per tale motivo, questa sostanza viene data al paziente per via intratecale. La tossina blocca il dolore, ma la sua azione sul sistema nervoso centrale può ancora produrre effetti collaterali come stordimento, nausea, mal di testa e confusione. Lo ziconotide, pertanto, viene dato per lo più a pazienti con un tumore all'ultimo stadio che non riescono a trovare sollievo in nessun altro modo.

Sono allo studio anche farmaci che agiscono sui recettori per i cannabinoidi. Sembra che questi agenti possano attenuare il dolore in diversi modi: per esempio interferendo con la trasmissione dei segnali tra i nocicettori e le loro cellule bersaglio, o riducendo l'attività delle cellule infiammatorie.

BERSAGLI FARMACOLOGICI ALLA PERIFERIA

Le minuscole ramificazioni dei nocicettori che innervano la pelle e gli organi interni hanno recettori che rilevano gli stimoli dannosi. Tra questi stimoli ci sono la capsaicina, la molecola piccante dei peperoncini, il calore intenso, o sostanze (come le prostaglandine) che vengono liberate dalle cellule infiammatorie in risposta a una ferita.
La percezione di questi stimoli da parte dei recettori fa sì che alcuni di essi permettano l'ingresso nelle cellule di ioni sodio e calcio. Questo flusso entrante, o l'attivazione di altri recettori, induce i nocicettori a emettere segnali dolorifici, e può spingere le cellule a rispondere a stimoli innocui. Anche la propagazione dei segnali richiede l'attivazione da parte dei canali ionici voltaggio-dipendenti. L'inibizione dell'attività delle molecole rilevatrici o dei canali dovrebbe avere effetti terapeutici.

Sigillare le entrate

Alcuni ricercatori stanno cercando di impedire che i neuroni spinali rispondano ai neurotrasmettitori rilasciati dai nocicettori, e in particolare al glutammato, che è il principale veicolo del messaggio dolorifico. Il glutammato attiva recettori nel corno dorsale del midollo spinale. Fra questi, i recettori della classe NMDA hanno un ruolo attivo nel processo di sensibilizzazione a livello centrale, e ciò ne fa un ovvio bersaglio per nuovi analgesici.

Ogni neurone è dotato di alcuni tipi di recettori NMDA, perciò se inibissimo simultaneamente questi recettori produrremmo effetti catastrofici come perdita di memoria, colpi apoplettici e paralisi. Per evitare simili reazioni, si sta tentando di inattivare i recettori presenti principalmente nel corno dorsale. I composti che si legano a una variante contenente la subunità N2RB hanno prodotto risultati incoraggianti. Topi in cui l'inibitore NR2B è stato somministrato nella spina dorsale, per esempio, hanno manifestato minor sensibilità rispetto agli animali non trattati.

Molti nocicettori rilasciano anche neurotrasmettitori peptidici, come la «sostanza P» e il peptide correlato al gene della calcitonina (CGRP). Poiché questi peptidi attivano i neuroni del midollo spinale che trasmettono gli stimoli dolorifici agendo su recettori distinti, potremmo aspettarci

LE NUOVE MOLECOLE DELLA GUERRA AL DOLORE

La tabella qui sotto elenca alcuni composti innovativi, attualmente in fase di sperimentazione. Non sono però menzionate nuove varianti di prodotti farmaceutici già esistenti, come gli inibitori degli oppiacei e gli inibitori della COX. I trial clinici procedono per fasi successive: la fase I si concentra sulla sicurezza, la fase II analizza l'efficacia con trial preliminari, mentre la fase III prevede test di più ampio respiro.

COMPOSTO (AZIENDA PRODUTTRICE)	MECCANISMO DI AZIONE	FASE DELLA SPERIMENTAZIONE	CASA FARMACEUTICA INTERESSATA
AMG-517 (Amgen)	Blocca il recettore per la capsaicina	Fase 1	GlaxoSmithKline; Neurogen
EVT-101 (Evotec)	Blocca i recettori NMDA che recano la subunità NR2B	Fase 1	Roche; Merck
Icatibant (Sanofi-Aventis)	Blocca un recettore per la bradichinina	Fase 2	Merck
NGX-4010 (NeurogesX)	Sovrastimola i recettori per la capsaicina	Fase 3	
NMED-160 (Neuromed Pharmaceuticals)	Blocca i canali del calcio di tipo N	Fase 2	
Ralfinamide (Newron Pharmaceuticals)	Blocca i canali del sodio	Fase 2	
RN624 (Rinat Neuroscience)	Arresta il fattore di crescita nervoso impedendogli di stimolare i nocicettori	Fase 2	Amgen
SAB-378 (Novartis)	Attiva un recettore per i cannabinoidi	Fase 2	GW Pharmaceuticals; GlaxoSmithKline

Fonte: Franz F. Hefti Rinat Neuroscience Corporation

che i farmaci che impediscono l'interazione con tali recettori siano efficaci. Ma nei trial clinici sul dolore il blocco selettivo del recettore usato dalla sostanza P (neurochinina-1, o NK-1) non ha avuto successo, forse perché è insufficiente. Non si sa ancora se ridurre l'attività del CGRP nel midollo spinale potrebbe eliminare il dolore, anche se si stanno sviluppando antagonisti che puntano ad alleviare il tormento delle emicranie interferendo con il rilascio del CGRP sui vasi sanguigni sulla superficie del cervello.

Eliminare i messaggeri?

Se i tentativi di modulare la trasmissione dei messaggi di dolore fallissero, si potrebbe pensare di eliminare il messaggero. La recisione dei nervi nocicettivi ha però effetti negativi, perché le lesioni nervose possono favorire l'insorgenza di un dolore ancora più ostinato e persistente.

Un tempo era pratica comune recidere i circuiti del midollo spinale che trasportano le informazioni al cervello. Oggi però questo intervento è ristretto ai malati terminali di tumore afflitti da dolore incessante refrattario a qualsiasi genere di trattamento. Il problema di questa procedura è che il chirurgo non può recidere selettivamente i circuiti del dolore. Una possibile soluzione, che ha prodotto risultati positivi sugli animali, è una terapia che elimina un sottogruppo di neuroni del midollo spinale che ricevono i segnali dai nocicettori. Questa terapia abbina una tossina chiamata saporina con la sostanza P. La sostanza P che è parte del coniugato si lega ai recettori NK-1 determinando l'internalizzazione dell'intero composto, dopodichè la saporina si libera e può uccidere il neurone. Siccome il complesso coniugato può entrare solo nelle cellule che hanno un recettore NK-1, i ricercatori sperano che gli effetti collaterali siano limitati.

Ma l'eliminazione dei neuroni nel midollo spinale deve considerata un'ultima spiaggia: i neuroni del sistema nervoso centrale non ricrescono, e quindi i cambiamenti provocati da questa procedura sono permanenti. Non accade altrettanto nel sistema nervoso periferico, dove invece le fibre recise si possono rigenerare. Idealmente, terapie che recidono le ramificazioni dei nocicettori per il rilevamento dei segnali – come quelle a base di dosi elevate di capsaicina – dovrebbero arrestare il dolore, ma consentire infine alle ramificazioni di ricrescere ripristinando la percezionefisiologicadel dolore nel frammento di tessuto innervato in precedenza.

Puntare ai neuroni potrebbe non essere il solo modo per eliminare il dolore. Diversi studi indicano che la glia, cioè le cellule che alimentano i neuroni nel sistema nervoso centrale, si attiva in risposta a un danno che interessa i nervi periferici. La glia migra nelle regioni del corno dorsale, e qui scarica sostanze che inducono i terminali dei nocicettori a rilasciare neurotrasmettitori nel midollo, facendo persistere il segnale di dolore. Anche alcune di queste sostanze, tra cui i fattori di crescita e le citochine, rendono eccitabili i neuroni del corno dorsale, cosicché i composti che bloccano questa iperattività dovrebbero contribuire a ridurre l'ipersensibilità. Diversi gruppi stanno lavorando per identificare – e scoprire il modo di inibire – le molecole che attivano la glia quando i nervi vengono danneggiati.

Stranamente, fra le sostanze chiave rilasciate dalla glia attivata nel midollo spinale ci sono le prostaglandine. In questa sede, esse aumentano il dolore, bloccando i recettori per la glicina situati sul corno dorsale dei neuroni. Di solito la glicina, un neurotrasmettitore inibitorio, calma questi neuroni. I FANS, dunque, possono agire anche inibendo gli enzimi COX nella glia. In questo caso, la somministrazione diretta di inibitori della COX nel midollo spinale potrebbe ridurre gli effetti collaterali causati dall'immissione sistemica. Ma anche un farmaco grado di aumentare l'attività del recettore per la glicina potrebbe aiutare a comprimere la trasmissione dei segnali verso il cervello.

PER APPROFONDIRE

BASBAUM A.I. e JESSEL T., *The Perception of Pain*, in *Principles of Neural Science*, Eric R. Kandel e altri (*a cura*), McGraw-Hill, 2000.

BASBAUM A.I. e JULIUS D., *Molecular Mechanisms of Nociception*, in «Nature», Vol. 413, pp. 203-210, 13 settembre 2001.

MCMAHON S.B., CAFFERTY W.B. e MARCHAND F., *Immune and Glial Cell Factors as Pain Mediators and Modulators*, in «Experimental Neurology», Vol. 192, n. 2, pp. 444-462, 2005.

Pain Collection, in «Nature Reviews Neuroscience», luglio 2005. Disponibile on line all'indirizzo: www.nature.com/nrn/focus/pain.

BERSAGLI FARMACOLOGICI NEL MIDOLLO SPINALE

Affinché i messaggi dolorifici provenienti dai nocicettori siano trasmessi al cervello dai neuroni del midollo spinale, è necessario che i nocicettori rilascino nel corno dorsale del midollo segnali chimici come il glutammato e la sostanza P. Queste molecole, poi, devono essere percepite da recettori specifici situati sui neuroni del corno dorsale. Inoltre, perché i nocicettori possano liberare le loro molecole segnale, è necessario che funzionino i canali del calcio sui terminali dei nocicettori, e che, viceversa, non funzionino i recettori per gli oppiacei. La morfina e gli oppiacei della stessa famiglia, che al momento sono tra le sostanze più efficaci, agiscono attivando i recettori per gli oppiacei. Tuttavia i derivati dell'oppio possono indurre effetti collaterali inaccettabili, ed è per questo motivo che i ricercatori stanno cercando nuove sostanze in grado di agire su bersagli diversi nel midollo spinale.

Amadeo Bachar

Un problema di percezione

Tra gli approcci che abbiamo illustrato, quelli più interessanti non modificano le normali sensazioni, ma riducono o eliminano l'ipersensibilità caratteristica dei dolori infiammatori e neuropatici, e hanno effetti collaterali accettabili. Ma queste terapie aiuteranno davvero i pazienti? E agiranno sul dolore a tutti i livelli? A queste domande non c'è ancora risposta.

Un approccio che vale la pena approfondire è l'uso di terapie comportamentali, cioè terapie non farmacologiche, in particolare per il dolore associato a condizioni come la fibromialgia e la sindrome da colon irritabile, per le quali non è stata ancora definita una causa organica. Una decina d'anni fa alcuni ricercatori della McGill University hanno dimostrato che l'ipnosi poteva alterare l'attività cerebrale di un soggetto e la sua percezione del dolore. Gli scienziati avevano ipnotizzato dei volontari e avevano suggerito loro che il bagno d'acqua calda nel quale avevano immerso le mani fosse più, o meno, sgradevole di quanto era in realtà.

Usando la PET per monitorare l'attività cerebrale, si è scoperto che la corteccia somatosensoriale era altrettanto attiva in entrambe le situazioni. E una seconda regione cerebrale, la corteccia del cingolo, era più attiva quando i soggetti credevano che lo stimolo fosse più sgradevole: l'ipnosi aveva cambiato il modo in cui percepivano le sensazioni. Acquisendo altre informazioni sul modo in cui il cervello modula l'esperienza del dolore, si potrebbero sviluppare migliori terapie cognitive per moderare la percezione del dolore.

Molte strade sono aperte. Speriamo che continuando a svolgere ricerche sui meccanismi del dolore si possa giungere a terapie sicure ed efficaci.

FEMMINA TIGRE. *Aedes albopictus* ha dimensioni simili a quelle di una normale zanzara, circa un centimetro, ma può essere riconosciuta facilmente a occhio nudo per la caratteristica colorazione nera a macchie bianche; in particolare, permettono di identificare la specie la presenza di una sottile striatura bianca sul dorso e i piccoli anelli bianchi sulle zampe.

Sinclair Stammers/SPL/Grazia Neri

Zanzara tigre: un nemico sottovalutato

Considerata da molti soprattutto come fonte di fastidio, *Aedes albopictus* è anche un insetto potenzialmente pericoloso per l'uomo

di Francesco Severini, Marco Di Luca, Luciano Toma e Roberto Romi

Aedes albopictus, la famigerata zanzara tigre, è ormai da 15 anni ospite indesiderato dell'Italia. In molte zone del nostro paese la sua presenza è anzi così radicata da poterla considerare un elemento integrante della nostra fauna entomologica. Fino al secondo dopoguerra, l'areale di distribuzione della zanzara tigre si estendeva dal Subcontinente indiano al Giappone, includendo il Sudest asiatico, le isole dell'Oceano Indiano e il Madagascar. L'insetto era legato ad ambienti di foresta, dove i focolai di sviluppo erano le cavità degli alberi, i bambù spezzati e le ascelle fogliari di una famiglia di piante succulente (le bromeliacee, si veda la foto a p. 104) e di altre piante che si riempiono d'acqua con i violenti temporali tipici dei paesi tropicali e subtropicali.

L'espansione delle attività umane a ridosso delle foreste tropicali ha indotto la zanzara tigre, dotata di grande plasticità ecologica, ad adattarsi ad altri focolai tipici degli ambienti antropizzati: in genere si tratta di contenitori medio-piccoli, con minime quantità d'acqua, presenti sia nelle aree commerciali e industriali dei centri abitati sia nell'ambiente domestico. La sua adattabilità ha permesso alla specie di sfruttare come focolai larvali anche i copertoni d'auto abbandonati all'aperto, tanto che il commercio di pneumatici da rigenerare è stato il principale mezzo di diffusione dell'insetto al di fuori dell'areale d'origine. Oggi *Aedes albopictus* è considerata un *container breeding mosquito* (zanzara che si riproduce in contenitori), tipica delle aree densamente popolate.

Con l'aumentare degli scambi commerciali, la zanzara tigre ha colonizzato il continente americano, dove si è rapidamente diffusa in tutti gli Stati degli Stati Uniti a est delle Montagne Rocciose e, seguendo la costa atlantica, in America Latina fino al Brasile. Quindi è giunta in Europa, insediandosi prima in Albania, e poi si è diffusa in Italia (nel 1990), in Francia (nel 2000), e recentemente in Olanda, in Svizzera, in alcuni Stati balcanici (Bosnia, Slovenia, Croazia, Serbia, Montenegro), in Spagna e in Grecia, dove sembra essere ancora presente con pochi focolai circoscritti.

Le caratteristiche della specie

Il ciclo biologico della zanzara tigre è simile a quello delle zanzare italiane appartenenti allo stesso genere e alle nostre latitudini; a seconda delle condizioni climatiche, può durare dai dieci giorni alle tre settimane (*si veda l'illustrazione a p. 105*).

Originaria di ambienti tropicali caratterizzati da temperature e umidità nettamente superiori alle nostre, *Aedes albopictus* è riuscita in breve tempo a colonizzare ecosistemi completamente diversi. Molti insetti tropicali sono incapaci di sopravvivere ai rigidi mesi invernali delle nostre latitudini, ma la zanzara tigre è riuscita a superare questo ostacolo grazie all'introduzione nel nostro paese di popolazioni provenienti dai limiti settentrionali dell'areale d'origine. Queste zanzare sono infatti dotate di un «orologio biologico» che induce la deposizione di uova particolarmente resistenti al freddo (le cosiddette uova diapausanti), quando le

giornate iniziano ad accorciarsi e la temperatura a diminuire. Le uova diapausanti si schiudono solo quando le condizioni climatiche, in particolare il numero di ore di luce, tornano a essere favorevoli allo sviluppo della nuova generazione.

La distribuzione in Italia

Dopo il suo arrivo in Italia, all'inizio degli anni novanta, *Aedes albopictus* si è diffusa in modo rapido dal principale sito d'ingresso, il Veneto, al resto del paese, formando nuove colonie distribuite in modo discontinuo sul territorio. La specie si è radicata saldamente nelle zone in cui i fattori ambientali e climatici sono stati più favorevoli, condizioni presenti nelle aree di pianura o di bassa collina con clima caldo-umido, dove l'elevata umidità nei mesi estivi è assicurata da precipitazioni abbondanti e dalla ricca vegetazione. Nella penisola la specie è presente in quasi tutte le Regioni, con focolai particolarmente importanti nella Pianura Padana, nelle zone collinari delle regioni nord-orientali, nella parte più settentrionale della costa adriatica e lungo tutta la costa tirrenica. Più di recente sono stati individuati i primi focolai siciliani, nelle province di Messina e Palermo.

In generale nelle regioni centro-settentrionali, dove la quantità annuale di pioggia è superiore ai 600 millimetri, le condizioni ambientali consentono uno sviluppo abnorme della specie, mentre in quelle meridionali, caratterizzate da minori quantità di precipitazioni, la zanzara tigre ha maggiori difficoltà a stabilire focolai permanenti o comunque a riprodursi in modo considerevole.

Ma nei grandi agglomerati urbani sono altri i fattori che svolgono un ruolo importante nel determinare le condizioni adatte allo sviluppo. In queste aree la presenza di attività industriali e commerciali (per esempio vivai, depositi di copertoni, rottamazioni d'auto) favorisce l'insediamento della specie; in seguito, la disponibilità di raccolte d'acqua adatte allo sviluppo larvale in orti, giardini e terrazzi permette alla zanzara di diffondersi rapidamente nel resto dell'area urbana anche in assenza di precipitazioni atmosferiche.

I tombini per lo smaltimento delle acque piovane sono una categoria importante di focolai larvali, perché al loro interno si trovano piccoli invasi (le caditoie) che trattengono una certa quantità d'acqua ricca in contenuto organico, ideale per lo sviluppo delle larve di *Aedes albopictus*. In Italia, le caditoie rappresentano il principale, se non l'unico, bersaglio degli interventi di disinfestazione condotti sul suolo pubblico. Ma è difficile stabilire quanto questi focolai influiscano sulla produzione totale di larve, perché spesso buona parte di essi è asciutta durante i mesi estivi.

Come difendersi

La lotta contro la zanzara tigre va condotta adottando misure diverse, in stretta collaborazione tra le strutture pubbliche preposte al controllo (aziende sanitarie o Comuni) e i cittadini, con l'obiettivo di ridurre la densità della specie a livelli tollerabili.

Le strutture pubbliche dovrebbero monitorare attivamente la presenza e la densità della zanzara, sia per rilevarne precocemente l'ingresso in aree ancora indenni, e dunque avere qualche possibilità di eliminare il focolaio, sia per guidare stagionalmente gli interventi di contrasto nelle aree già infestate, prima che la specie raggiunga densità preoccupanti.

I cittadini dovrebbero eliminare o almeno svuotare periodicamente tutti i potenziali focolai presenti nell'ambiente domestico (per esempio sottovasi, secchi, bidoni, annaffiatoi). I focolai che non possono essere rimossi (per esempio i chiusini nei cortili) andrebbero trattati periodicamente con prodotti specifici, i «larvicidi», da poco immessi sul mercato anche in confezioni per uso non professionale. Nei casi in cui è possibile, devono essere privilegiati prodotti biologici a base di tossine e spore del batterio *Bacillus thuringiensis*, innocui per l'uomo e per gli animali domestici.

È possibile proteggere dalle punture le parti scoperte del corpo con l'uso di repellenti da applicare sulla cute. Ma visto che sono veramente efficaci solo i prodotti sintetizzati in laboratorio, potenzialmente tossici anche per l'uomo, è bene impiegarli solo quando è strettamente necessario. Per la protezione delle abitazioni è preferibile usare le zanzariere al posto dei comuni elettroemanatori per le piastrine antizanzara (i fornelletti) e degli spray, che rilasciano anch'essi sostanze potenzialmente tossiche.

Ma per evitare di favorire un'ulteriore diffusione della zanzara tigre e di rendere il problema ancora più acuto, il controllo della zanzara tigre deve essere guidato da un capillare monitoraggio del territorio ed essere sostenuto da un'approfondita attività di studio e di ricerca.

FOCOLAIO LARVALE DELLA ZANZARA TIGRE in un ambiente forestale del Sudest asiatico. Le larve si sviluppano nelle piccole raccolte d'acqua piovana presenti nelle ascelle fogliari di alcune piante grasse, le bromeliacee.

GLI AUTORI

FRANCESCO SEVERINI, MARCO DI LUCA, LUCIANO TOMA sono biologi e naturalisti dell'unità, coordinata da ROBERTO ROMI, che si occupa di entomologia medica presso il Dipartimento di malattie infettive, parassitarie e immunomediate dell'Istituto superiore di Sanità.

UN PROBLEMA SANITARIO

Le zanzare, come molti altri insetti, occupano una posizione fondamentale all'interno della catena alimentare perché molti animali (pesci, anfibi, rettili, uccelli e mammiferi) sono insettivori. Ma, oltre a svolgere un importante ruolo ecologico, moltissime specie di zanzare sono vettori di agenti patogeni responsabili di infezioni umane e animali in ogni parte del mondo.

Tra le numerose malattie ricordiamo la malaria (malattia causata dai protozoi del genere *Plasmodium* e trasmessa da zanzare del genere *Anopheles*), che ogni anno causa migliaia di vittime, la dengue, la febbre gialla, la sindrome da West Nile e altre provocate da arbovirus (gruppo di virus trasmessi principalmente da zanzare dei generi *Aedes*, *Ochlerotatus* e *Culex*). Nella fauna indigena italiana sono presenti i potenziali vettori di malaria e arbovirus, ma mancano i serbatoi umani e animali necessari per trasmettere il contagio.

Non vanno inoltre sottovalutate le piccole «emergenze sanitarie» dovute alla presenza di densità elevate di zanzare in molte zone del nostro paese. In seguito alle punture, persone particolarmente sensibili possono manifestare una reazione allergica localizzata con formazione di piccole lesioni cutanee dolorose e spesso edematose o emorragiche. Questi effetti sono maggiormente visibili su bambini e anziani, e spesso richiedono un intervento medico.

ZANZARA TIGRE: CHE COSA SI RISCHIA?

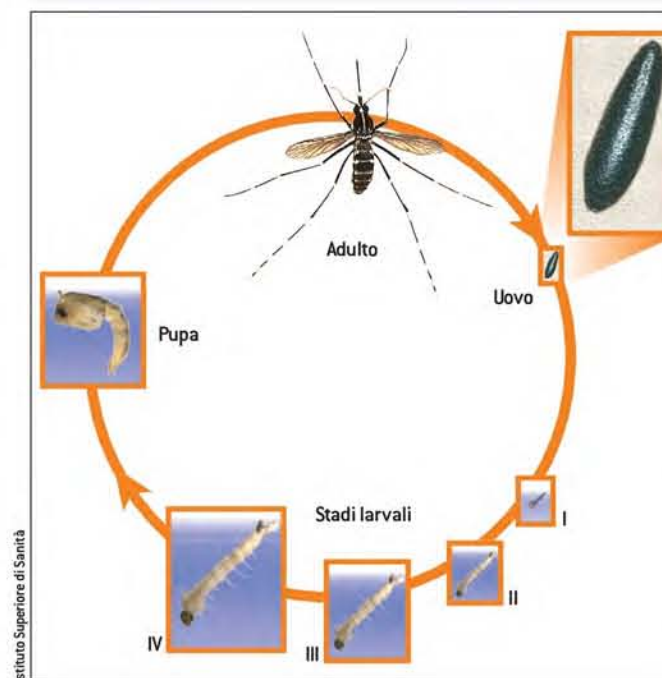
Nei paesi d'origine, la zanzara tigre è un efficiente vettore di virus che sono assenti in Italia e in tutto il bacino del Mediterraneo. In un futuro, però, i cambiamenti climatici potrebbero modificare lo scenario, e consentire l'introduzione di questi virus esotici anche nei paesi a clima temperato.
Gli esperimenti effettuati in laboratorio hanno dimostrato che la zanzara tigre è in grado di trasmettere oltre 20 virus, ma si valuta che il rischio più concreto sia quello legato alla trasmissione del virus della dengue.

La dengue è una malattia endemica in tutta la fascia tropicale dell'Asia, del Pacifico meridionale e del continente americano, trasmessa all'uomo da zanzare appartenenti al genere *Aedes*, tra cui *Aedes aegypti* (lo stesso vettore della febbre gialla), e la stessa *Aedes albopictus*.

La malattia si presenta in due forme diverse, e in genere segue un andamento epidemico. La prima è una sindrome febbrile a evoluzione benigna, diffusa in tutta l'area dove la patologia è endemica, detta anche «febbre spacca-ossa» per via dei dolori muscolari e articolari che la caratterizzano.

La seconda forma si presenta come una sindrome emorragica spesso a esito fatale, ed è diffusa nelle aree urbane del Sudest asiatico e dell'America centrale. In Italia la malattia è solo d'importazione (circa 50 casi l'anno).

Di recente esemplari di zanzara tigre infetti con il virus *Chikungunya*, agente di una malattia simile alla dengue, sono stati ritrovati in alcune isole dell'Oceano Indiano, con numerosi casi di importazione nelle aree metropolitane della Francia.

Le filariosi sono malattie causate da vermi parassiti appartenenti alla superfamiglia *Filarioidea*. Le filarie più patogene per l'uomo sono presenti solamente in paesi tropicali e subtropicali, mentre in Italia quelle più comuni sono parassiti del cane e del gatto (genere *Dirofilaria*). Questi parassiti, però, possono essere trasmessi accidentalmente all'uomo da zanzare di generi diversi, soprattutto in zone rurali. Di recente nel Veneto e nel Lazio sono stati catturati esemplari di zanzara tigre infetti con *Dirofilaria immitis* e *Dirofilaria repens*. La zanzara tigre potrebbe dunque amplificare la trasmissione di questi patogeni anche in aree densamente popolate.

CICLO BIOLOGICO. Le uova, deposte su un substrato asciutto poco sopra il livello dell'acqua, si schiudono circa 1-2 giorni dopo essere state sommerse. Da ognuna fuoriesce una larva di primo stadio che con quattro mute raggiunge lo stadio di pupa e da quest'ultimo, dopo circa 48 ore, arriva all'esemplare adulto (la cosiddetta immagine). Tutti gli stadi precedenti l'immagine sono acquatici. La femmina adulta si accoppia un'unica volta 24-48 ore dopo la nascita, poi va in cerca di un ospite sul quale effettuare il pasto di sangue e due giorni dopo depone le uova. In estate, con temperature medie superiori a 25 gradi, l'intero ciclo biologico si può completare in 8-10 giorni.

PER APPROFONDIRE

DI LUCA M., SEVERINI F., TOMA L. e ROMI R., *Zanzara Tigre: un raffinato esempio di plasticità ecologica*, su «Biologi Italiani», n. 6, pp. 36-43, 2003.

ROMI R., *Aedes albopictus in Italia: un problema sanitario sottovalutato*, su «Annali dell'Istituto superiore di Sanità», Vol. 37, n. 2, pp. 241-248, 2001.

SABATINI A., RAINERI V., TROVATO G. e COLUZZI M., *Aedes albopictus in Italia e possibile diffusione nell'area mediterranea*, su «Parassitologia», Vol. 32, pp. 301-304, 1990.

Slim Films

Il sogno sfuggente della traduzione automatica

di Gary Stix

I metodi statistici riusciranno finalmente a rendere veloci (e soprattutto affidabili) le traduzioni eseguite dai computer?

La leggenda sodio Nepal Asia: il leone, la strega, lo spirito malvagio dell'armadio «manca già» lo spirito malvagio si astiene dalla trilogia «ricca di poetica e di gusto artistico, non ha nemmeno lasciato» Harley Baud «la serie di racconti ha il piacere infinito della corrente nascosta che è turbolenta.

Questo incomprensibile groviglio di parole è come suonerebbe in italiano la traduzione eseguita da Babelfish, il traduttore automatico on line di Altavista, di un passo in cinese tratto dalla pagina web del «China Post» di Taiwan. Il testo originale diceva invece:

> *Le Cronache di Narnia* non presenta le poetiche visioni della trilogia di *Il Signore degli Anelli* e nemmeno gli oscuri misteri che rendono irresistibile e affascinante la serie di *Harry Potter*.

Questo esempio mostra come la traduzione automatica sia uno dei settori dell'intelligenza artificiale dove la sfida è ancora aperta; è sufficiente un nome proprio o una frase un po' più articolata del normale per mettere fuori strada il software. Negli ultimi anni, però, i ricercatori hanno seguito un approccio diverso, che ha riacceso l'interesse per la traduzione automatica: i programmi *brute-force*, che valutano il grado di corrispondenza tra parole o frasi di due lingue differenti, stanno finalmente avvicinando qualitativamente la traduzione automatica a quella professionale, o almeno così sostengono gli sviluppatori di software.

Più difficile degli scacchi

Ormai la potenza sempre maggiore degli algoritmi hardware e software ha permesso a una macchina di trionfare addirittura su un grande maestro di scacchi (è il caso del supercomputer IBM Deep Blue che nel 1997 sconfisse Garry Kasparov). Ma la traduzione automatica, in oltre cinquant'anni di ricerche, ha compiuto solo pochi, incerti passi verso risultati che eguagliano le capacità di un essere umano, e ad alcuni critici quest'affermazione potrebbe sembrare eccessivamente generosa.

Nel 1954 la IBM, assieme all'Università di Georgetown, tenne la prima dimostrazione di traduzione automatica dal russo in inglese di oltre 60 frasi, e nel comunicato stampa diffuso l'8 gennaio dello stesso anno annunciò: «Oggi, per la prima volta, il russo è stato tradotto in inglese da un "cervello" elettronico». I militari impiegati nella difesa degli

TRADUZIONE AUTOMATICA BASATA SUL METODO STATISTICO

INSERIMENTO DI TESTI GIÀ TRADOTTI
I testi già tradotti, presi da diverse fonti, costituiscono la base della traduzione automatica.

Archivi di testi già tradotti
Dizionari e glossari
Internet
Traduttori professionisti

È stato dimostrato che i metodi statistici sono più efficaci dei sistemi basati sulle regole grammaticali inserite da traduttori professionisti.
Le nuove metodologie si avvalgono della capacità dei computer di analizzare i testi già tradotti per determinare il grado di probabilità che a una parola nella lingua di partenza corrisponda una specifica parola nella lingua di arrivo.

ELABORAZIONE DEI DATI I testi sono scanditi, allineati e formattati.

Que hambre tengo yo.
Rielaborazione
que hambre tengo yo

CREAZIONE DELLE CORRISPONDENZE TRA LE FRASI
Un modello di traduzione sceglie frasi di due o tre parole dal testo di partenza (in questo caso in spagnolo) e le associa a frasi nella lingua di arrivo (in inglese).

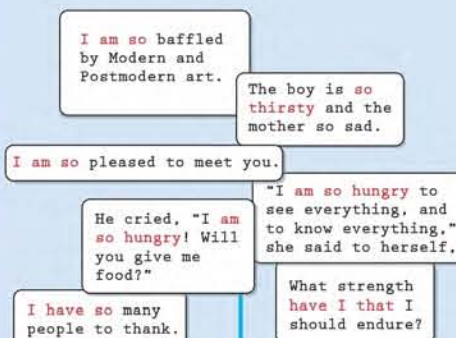
Lingua di partenza: *spagnolo*
Este guiso tradicional se ennoblece con el bogavante, la viera y el rodaballo.
Que hambre tengo yo.

Lingua di arrivo: *inglese*
This traditional stew is refined with scallops, lobster and turbot.
I am so hungry

MODELLI DI TRADUZIONE
Usando la statistica per calcolare la frequenza e la collocazione delle parole nelle frasi nelle due lingue, il modello costruisce uno schema dell'ordine delle parole nella frase. Inoltre il modello sfrutta anche altre tecniche come la riduzione di diversi termini spagnoli in una singola parola (non mostrato).

Que hambre tengo yo
I
am
so
hungry

MODELLI LINGUISTICI
Sulla base di precedenti analisi statistiche di testi scritti in inglese, un modello linguistico cerca di prevedere la traduzione più probabile di una frase o di una parola in un testo già tradotto; più alta è la frequenza di una frase, maggiore è la probabilità che si tratti della traduzione corretta.

I am so baffled by Modern and Postmodern art.
The boy is so thirsty and the mother so sad.
I am so pleased to meet you.
"I am so hungry to see everything, and to know everything," she said to herself.
He cried, "I am so hungry! Will you give me food?"
I have so many people to thank.
What strength have I that I should endure?

I am so > Have I that
I am so > I have so
So thirsty > Thirsty
Am so hungry > What hunger have

DECODIFICA Quando si inserisce una nuova frase – che si differenzia parzialmente o interamente da quelle già inserite in precedenza (per esempio, usando *sed*, sete, al posto di *hambre*, fame) – il decoder elabora diverse possibili traduzioni e sceglie quella più probabile.

TESTO INSERITO
Que sed tengo yo.

DECODIFICA

I am so thirsty	P = 0,13
What thirst have I	P = 0,09
Have I what thirst	P = 0,07
Thirsty I am so	P = 0,00

TRADUZIONE
I am so thirsty.

Lucy Reading-Ikkanda; Kevin Knight, *Language Weaver*, e Philipp Koehn, *Università di Edimburgo (fonti)*

Stati Uniti e gli informatici si aspettavano che entro cinque anni la traduzione automatica sarebbe diventata una realtà quotidiana. Ma si sbagliavano.

Nel 1966, l'Automatic Language Processing Advisory Committee (ALPAC), una commissione finanziata dal Governo degli Stati Uniti, pubblicò un rapporto in cui si affermava che la traduzione professionale, ovvero quella umana, era più precisa, più veloce e costava circa la metà della traduzione automatica. La relazione concludeva: «Non si prevedono risultati utili, imminenti o futuri, nel campo della traduzione automatica».

I finanziamenti furono sospesi e nei decenni a seguire alla ricerca furono destinata solamente somme modeste. Verso la fine degli anni sessanta la U.S. Air Force, l'aeronautica militare degli Stati Uniti, per far fronte alla grande quantità di documenti da tradurre dal russo in inglese finanziò una piccola azienda, ideatrice di un traduttore automatico chiamato Systran. (La versione di questo software disponibile su Internet ha elaborato l'irripetibile traduzione con cui si apre questo articolo.)

Systran è un sistema basato su regole, perché fa riferimento alle norme grammaticali sia della lingua di partenza sia della lingua di arrivo; così operava pure l'originale «cervello» dell'IBM, basandosi su sei regole fondamentali che governano sintassi e semantica. Per tradurre in inglese la parola russa *o*, per esempio, un computer IBM 701 poteva scegliere tra due opzioni: il corrispettivo inglese *about* (in italiano: su, circa) oppure *of* (di). Se *o* appariva dopo la parola *nauka* (scienza), una regola indicava al computer che la traduzione corretta di *o* era *of* e non *about*, per ottenere la traduzione inglese *science of* (scienza di) e non *science about* (scienza su).

Nonostante la Systran sia tra le aziende più importanti a livello mondiale in materia di traduzione automatica e possa contare su clienti del calibro di Google, Yahoo e Time Warner-AOL, nel 2004 il suo profitto è stato di circa 13 milioni di dollari, in un mercato mondiale per le traduzioni, di ogni tipo, stimato intorno ai dieci miliardi di dollari. «Siamo piccoli, ma siamo i più grandi», dichiara Dimitris Sabatakakis, presidente e direttore generale della società, che ha sede a Parigi.

Basta regole!

Affinché i sistemi basati su regole possano creare un testo nella lingua in cui si vuole tradurre, al momento della realizzazione del programma i linguisti che ci lavorano devono inserirvi minuziosamente enormi vocabolari, regole grammaticali, sintattiche e semantiche; i software in commercio, per esempio, contengono decine di migliaia di regole grammaticali e centinaia di migliaia di vocaboli.

A partire dalla seconda metà degli anni ottanta l'IBM ha messo in commercio un software per la traduzione dal francese in inglese, chiamato Candide, che non si serviva di regole grammaticali e sintattiche, ma faceva invece riferimento a testi già tradotti e associava le parole corrispondenti nelle due lingue (sistemi più recenti associano intere frasi). Il sistema, infine, sulla base del teorema di Bayes sul calcolo delle probabilità, valutava il grado di correttezza della traduzione inglese di un termine francese.

Un'altra ricerca, basata unicamente sull'analisi di testi inglesi, mirava a verificare se una parola era adeguata al contesto grammaticale nel quale era inserita. Il termine o la frase nella lingua in cui si voleva tradurre con il più alto grado di probabilità che fosse esatto poteva essere usato successivamente per «decodificare» altri testi; in questo modo, più parole potevano essere legate a costruire periodi complessi. In teoria, perciò, era sufficiente avere la conferma statistica della corrispondenza tra il termine francese per «tormenta di neve» (*poudrerie*) e il corrispettivo termine inglese *blowing snow* per considerare corretta la traduzione.

Ma l'IBM decise di abbandonare questo settore di ricerca. Alla fine degli anni novanta, infatti, un traduttore automatico poteva impiegare un giorno intero per tradurre una sola pagina. Proprio allora però qualcosa ha iniziato a muoversi. Con l'avvento di Internet, si è assistito a una rapida crescita del numero di grandi testi scritti in due lingue, e inoltre la domanda di traduzioni si è impennata al punto da non poter più essere soddisfatta da traduttori in carne e ossa.

Nel 1999 la National Science Foundation ha tenuto un seminario alla Johns Hopkins University. La riunione aveva come obiettivo lo studio di un pacchetto di strumenti software che potesse essere facilmente impiegato dalla comunità scientifica; questo progetto attirò l'attenzione degli esperti e diede nuova energia alle ricerche. Nel 2002 Kevin Knight, uno degli organizzatori del seminario, e Daniel Marcu, suo collega alla University of Southern California, hanno fondato Language Weaver, l'unica società di traduzione automatica a impiegare il metodo statistico; attualmente la Language Weaver garantisce di poter tradurre 5000 parole al minuto dall'inglese in arabo, farsi, francese, cinese e spagnolo, e viceversa.

I primi tentativi di una traduzione automatica di testi hanno avuto come scenario l'inizio della Guerra Fredda

Google è un vincente

L'estate scorsa il sistema sperimentale progettato per Google da Franz Och, anch'egli presente al seminario ed ex alunno della University of Southern California, ha sbaragliato concorrenti del calibro della IBM nel corso della competizione, organizzata dal National Institute of Standards and Technologies, che prevedeva la traduzione dall'arabo o dal cinese in inglese di 100 documenti presenti su Internet. Secondo Och il segreto per migliorare l'efficacia del software è inserire l'equivalente di un milione di libri. Och ha confrontato la traduzione automatica dal cinese all'inglese allora adottata da Google (Systran) con il software sperimentale basato sulla statistica messo a punto da lui:

Google/Systran: «Il dottore indica, lo splendente nucleo predispone recupera il circa un mese.»

Google/Och: «I dottori dicono che Akihito dovrà stare a riposo per circa un mese.»

L'euforia per il metodo statistico applicato alla traduzione automatica ha messo in allarme la Systran. «Quando si studia una lingua straniera, bisogna impararne anche le regole grammaticali», ha commentato Sabatakakis, specificando: «Non si impara una lingua con il metodo statistico». Systran usa la statistica, ma per creare sistemi all'interno di domini molto piccoli, per esempio, per la traduzione di brevetti.

Ma per Sabatakakis l'attuale impiego del metodo statistico sta andando ben oltre le semplici strategie di mercato. Nel dipartimento ricerca e sviluppo di Systran lavorano 50 persone, e tra queste anche dei linguisti. «La principale differenza tra noi e Google è che loro sostengono di non aver bisogno di persone madrelingua per sviluppare le applicazioni in cinese perché sono sufficienti la magia e la bellezza del progetto», afferma Sabatakakis. E aggiunge: «Tuttavia, se non si impiega personale cinese il programma potrebbe commettere errori enormi».

La distanza tra i due metodi ha cominciato a ridursi quando i ricercatori hanno iniziato a inserire nei programmi basati sul metodo statistico dei parametri che tengono conto anche della struttura sintattica della frase. In questo modo è possibile fare a meno dell'intervento di un traduttore «umano»: un modello sintattico è infatti in grado di valutare la possibilità che una frase in inglese del tipo «aggettivo-nome» venga riordinata nella traduzione in italiano. Kevin Knight, di Language Weaver, sostiene che considerare le frasi anziché le singole parole permette alle statistiche di tenere conto anche dell'aspetto semantico, evitando, per esempio, che il suo cognome venga tradotto in «cavaliere».

Alla Microsoft il gruppo di elaborazione dei linguaggi naturali è numeroso. Negli ultimi sei anni ha lavorato anche nel campo della traduzione automatica, inizialmente studiando i sistemi basati su regole e spostando poi progressivamente l'attenzione verso i sistemi statistici. Di recente, la Microsoft ha usato soprattutto i metodi statistici per la traduzione in 12 nuove lingue, tra cui russo, arabo e cinese, per il proprio servizio di assistenza clienti; le traduzioni, però, non sono state pubblicate. «Alcune parti sono senza dubbio imprecise; altre invece sono piuttosto buone», racconta Steve Richardson, ricercatore alla sezione di elaborazione dei linguaggi naturali. «La qualità dei metodi statistici è paragonabile, se non addirittura superiore, a quella dei sistemi basati su regole che utilizzavamo in passato».

Il succo del discorso

Alla luce dello sviluppo di queste nuove tecniche, nasce spontanea la domanda se mai un software riuscirà a sconfiggere un traduttore professionista al suo stesso gioco, come è accaduto con il computer campione di scacchi dell'IBM. Una macchina può generare qualcosa di più di un'idea approssimativa del contenuto di un testo scritto in una lingua straniera? Kevin Hendzel, portavoce dell'American Translator Association sostiene che l'entusiasmo di oggi serve solo a giustificare decenni di eclatanti trovate pubblicitarie, come la Fully Authomatized High-Quality Translation, la traduzione di alta qualità completamente automatica, e sottolinea che fare riferimento al senso generale di un testo può essere d'aiuto se ci si trova a scegliere tra numerosi scritti, almeno fino a quando non ci si rende conto dell'inaffidabilità di questo criterio, perché anche una traduzione approssimativa comporta dei rischi.

Hendzel porta l'esempio di una traduzione dall'arabo all'inglese in cui erano descritte due parti «*going at each other*», ovvero «che si stavano attaccando»; questo passaggio ha attirato l'attenzione dei controlli di sicurezza, che poi hanno scoperto che nel testo si parlava di calcio e non di un imminente attacco terroristico o di una guerra.

Keith Devlin, direttore esecutivo del Center for the Study of Language and Information della Stanford University, assicura che la traduzione automatica non eguaglierà mai quella umana. «L'impiego di metodi statistici, affiancati da processori più veloci e memorie più ampie, porterà a sistemi di traduzione sempre migliori e in grado di operare in maniera accettabile in molti contesti – dice Devlin – ma, a mio parere, non si raggiungerà mai il livello di scorrevolezza di una traduzione eseguita da un traduttore in carne e ossa».

Knight, il padre del metodo statistico, non è d'accordo, e sottolinea i risultati ottenuti nell'arco degli ultimi dieci anni. Non vede ancora un limite alla tecnologia che, secondo lui, un giorno raggiungerà livelli umani in tutti i campi; a eccezione, forse, della poesia. Knight ha dimostrato che le traduzioni di uno stesso testo, una eseguita da un traduttore automatico e l'altra da un traduttore umano, sono indistinguibili per persone chiamate a riconoscerle, e afferma: «Non prendiamoci in giro; ci sono errori anche nelle traduzioni dei professionisti. La differenza non è così grande come si crede.

Per provare che questa volta si tratta di qualcosa di più della solita trovata pubblicitaria, il metodo statistico, che ora guida il settore, deve riuscire a dimostrare che la traduzione automatica di alta qualità è possibile. Solo allora si potrà dire che la tecnologia avrà fatto seguire i fatti alle parole.

Nonostante le numerose innovazioni, ancora non è chiaro se i computer potranno eguagliare le prestazioni umane

PER APPROFONDIRE

HUTCHINS J., *The History of Machine Translation in a Nutshell*. Disponibile all'indirizzo web: http://ourworld.compuserve.com/homepages/WJHutchins/nutshell.htm.

KNIGHT K., *A Statistical MT Tutorial Workbook*. Disponibile sul sito web: http://www.isi.edu/natural-language/mt/wkbk.rtf.

BERGER A.L. e altri, *The Candide System for Machine Translation*. Disponibile sul sito web: http://acl.ldc.upenn.edu/H/H94/H94-1028.pdf.